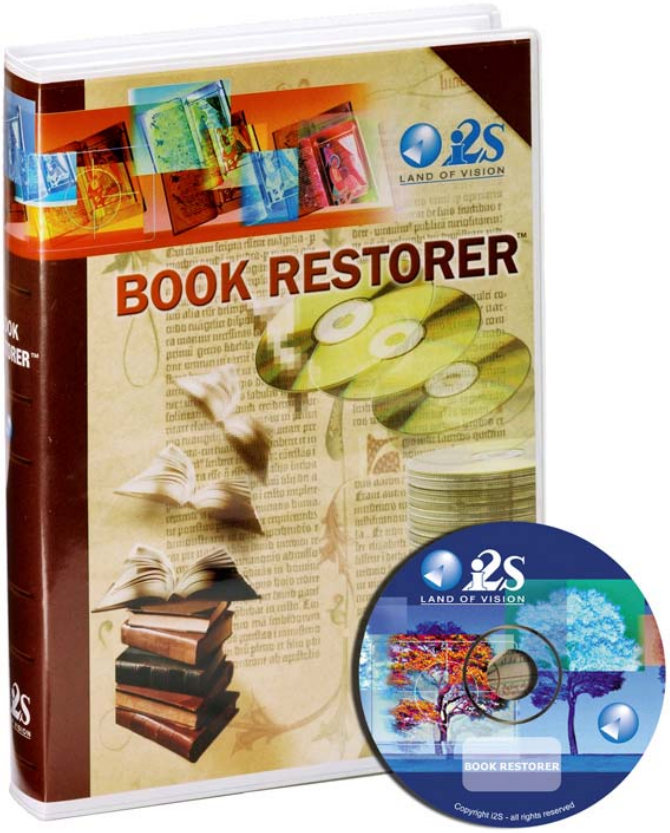
Алешин Д.А.

Основы работы в программе Book Restorer



16.08.2011

ОГЛАВЛЕНИЕ

1. ОБЩИЕ СВЕДЕНИЯ О ПРОГРАММЕ 3

1.1. Исполняемые модули программы 3

1.2. Цели и задачи, функциональность программы 4

1.3. Файлы и форматы Book Restorer 12

2. ИНТЕРФЕЙС ПРОГРАММЫ 14

2.1. Основные элементы интерфейса 14

2.2. Панель кнопок (линейка инструментов) и меню 15

2.3. Зона книги (дерево книги) 16

2.4. Окно процессов обработки ("реставрирование") 18

2.5. Панель (линейка) ретуширования 20

2.6. Лена прокрутки страниц, зона страницы и строка состояния 21

3. ИСПОЛЬЗОВАНИЕ BOOK RESTORER 22

3.1. Создание книги и «опубликование» обработанных страниц 22

3.2. Определение и редактирование зон в Book Restorer 27

3.3. Бинаризация 33

3.4. Геометрическая коррекция 35

3.5. Коррекция освещенности 42

3.6. К вопросу об использовании стилей 45

5. РАБОТА ДОПОЛНИТЕЛЬНЫХ МОДУЛЕЙ 50

5.1 BKR Automation - автоматизация обработки 50

5.2. BKR Watcher - контроль за папкой захвата 55

5.3. BKR Control - контроль качества 57

5.4. Автоматизация применения Book Restorer средствами OLE 60

Основы работы в программе Book Restorer

1. Общие сведения о программе

1.1. Исполняемые модули программы

Программа Book Restorer была разработана французской компанией i2S, которая изготавливает и продает планетарные сканеры и автоматические линии потокового сканирования. Book Restorer поэтому изначально предназначен для профессиональной промышленной работы с изображениями книжных страниц, создаваемыми этим оборудованием. На сайте разработчика: *http://www.i2s-digibook.com* (или второй, старый адрес *http://www.i2s-bookscanner.com*), а также его английского дилера *http://imageretrieval.co.uk* можно скачать триальную версию программы Book Restorer 4.2.1. Несмотря на различный размер инсталляционных пакетов, устанавливаются они одинаково.

На сайте другого, испанского дилера *http://www.vinfra.es* вплоть до 2011 года были доступны триальная версия 4.1, а также руководство пользователя. (Потом эти ссылки, увы, исчезли, но их копии сохранились на *www.archive.org*).

Программа Book Restorer разработана в соответствии с концепцией модульной архитектуры: отдельные элементы программы (исполняемые программы, процессы обработки) можно добавлять и убирать, не изменяя её общей работоспособности. Официальный релиз программы помимо основной программы для обработки изображений включает три дополнительных исполняемых модуля (файлы *.exe):

1) модуль автоматизации (BKR Automation) - пакетная обработка сформированных заданий, что ускоряет обработку однотипных документов и экономит время оператора, позволяя откладывать задания;

2) модуль "наблюдения" (BKR Watcher) - обработка изображений, как только они попадают в определенные папки;

3) модуль контроля качества (BKR Control) - создание отчета о процессе обработки. Осуществляется расчет коэффициента качества для каждого изображения и каждого процесса, что позволяет выявить изображения, которые позже следует проверить вручную.

Все процедуры обработки также сделаны в виде плагинов в основную программу (файлы *.dll). При наличии этих dll процессы доступны к запуску.

Такая модульная архитектура ведёт к тому, что распространяемые различные версии помимо совершенствования кода имеют различный состав дополнений к основной программе (указаны размеры в байтах):

версия	плагины процессов обработки		дополнительные модули	
4.1	18 файлов *.dll + BkrOCRPlugIn.dll	4.644.864 81.920	BKR Automation BKR Control BKR Watcher	475.136 430.080 413.696
4.2.1	18 файлов *.dll + BkrOpticalDistortion.dll	4.866.048 77.824	BKR Automation	471.040

Это означает, что из версии 4.1 в версию 4.2.1 можно скопировать пару исполняемых файлов, а также одну динамическую библиотеку, и собрать таким образом почти полноценную версию программы. Правда, имеющиеся в демо-версии 4.2.1. модули (BKR Automation) работают 100%, а добавленные "старые" у меня то запускается, то не запускаются. Хотя, опять-таки, в случае реальной необходимости их использования можно и версию 4.1 установить.

По такой же модульной системе, кстати, i2S сделала и справку к программе. Она разобрана на много мелких файлов и разложена по разным папкам. Хотя они и не всегда вызываются из основной программы, открыв вручную каждый из файлов *.hlp, можно прочитать всю необходимую информацию. В триальной версии 4.2.1. доступны стандартные файлы справки: *BookRestorer.hlp* - описание интерфейса программы; и *BKRCommonPlugin.hlp* - описание процедур обработки (расположен в папке Plugins).

На старом сайте *http://ww.i2s-bookscanner.com/* (см. его копию в *www.archive.org*) также был доступен файл справки, в котором дополнительно есть описание модуля Automation (*BkrAutomate.hlp)* и модуля Control (*BkrControl.hlp*). Там также можно найти файл справки к Book Restorer 3.1, в котором есть информация о функциях OLE Automation, встроенных в Book Restorer. Кроме того, по адресу *http://vinfra.es/soporte/bookrestorer_um.pdf*, как уже отмечалось, лежал "Book Restorer User Manual V4.0", а на сайте российского дилера - компании ЭЛАР - лежит аналогичный мануал на русском (адрес ссылки: *http://www.planscan.ru/po/rukovodstvo_po_book_restorer.doc*).

1.2. Цели и задачи, функциональность программы

Ключевыми особенностями программы, по утверждению разработчиков, ориентирующихся на ее промышленное использование, является:

1) имитирование структуры книги;

2) возможность обработки и вывода в потоковом режиме в качестве фонового процесса (при использовании модуля автоматизации).

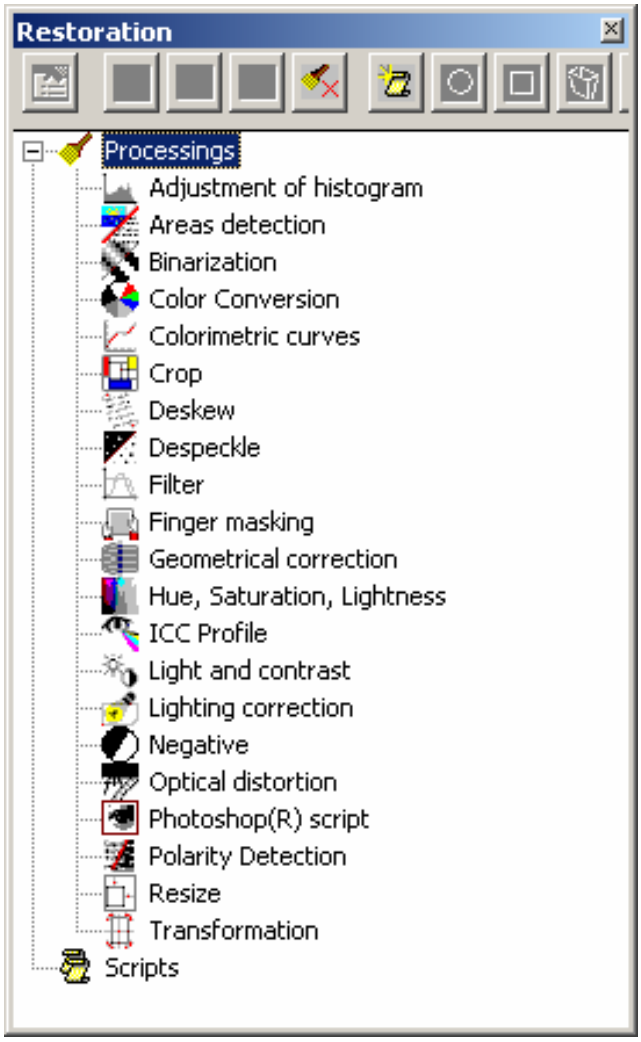
3) создание "скриптов" - наборов процедур для однотипной обработки одинаковых по качеству страниц.

Но на самом деле, Book Restorer является программой, которую необходимо иметь любому книголелу-любителю, который сканирует лишь пару книг в год, потому что в ней и только в ней есть эффективно работающая процедура исправления геометрических искажений. Такая процедура, к сожалению, отсутствует в какой-либо другой доступной программе. В Book Restorer также неплохо реализована адаптивная бинаризация (для которой, впрочем, есть аналог - Binarization Image Processor, но настолько неудобный, что на мой взгляд, лучше его даже и не пробовать). Определенный интерес представляет в Book Restorer и процедура коррекции освещенности, которая, хотя и есть в программах-аналогах, но работает по-другому.

Кроме того, важным отличием Book Restorer от другого программного обеспечения является его интуитивная понятность. Я бы вообще всем начинающим сканировщикам советовал попользоваться какое-то время Book Restorer, т.к. это дает понимание процесса.

В программе Book Restorer 4.2.1. существуют следующие процедуры обработки (по аналогии с другими программами обработки изображений их, наверное, можно назвать «инструментами реставрирования»):

- корректировка гистограммы
- определение зон
- бинаризация
- конвертация цветов
- колориметрическая кривая
- обрезка ("кроп")
- поворот
- удаление точек ("деспекл")
- фильтрация
- сокрытие пальцев
- геометрическая коррекция
- изменение цветовой модели
- профиль сканера
- освещенность и контраст
- изменение освещенности
- негатив (обращение цветов)
- оптическая коррекция
- скрипты фотошопа
- полярность
- изменение размера
- трансформация

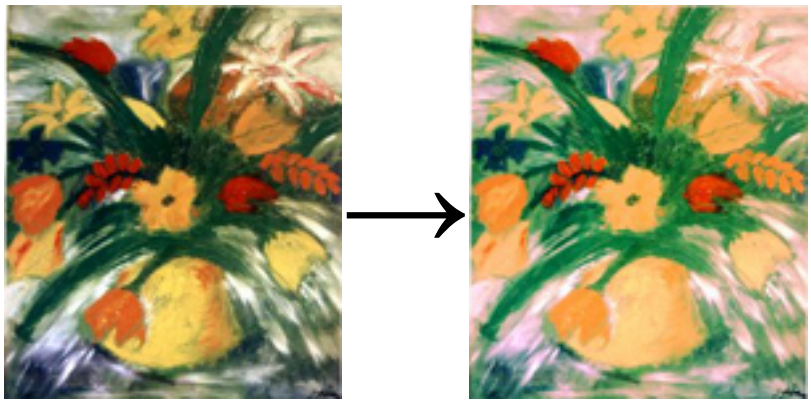


Еще один процесс обработки существовал в Book Restorer 4.1 - это OCR Module - процесс распознавания текста и создания текстовых блоков.

Рассмотрим результаты применения этих процессов обработки:

1. Adjustment of histogram

Корректировка гистограммы перераспределяет цвета по всем возможным значениям в соответствии с заданными параметром (аналог инструмента "уровни" в Photoshop).



2. Areas detection

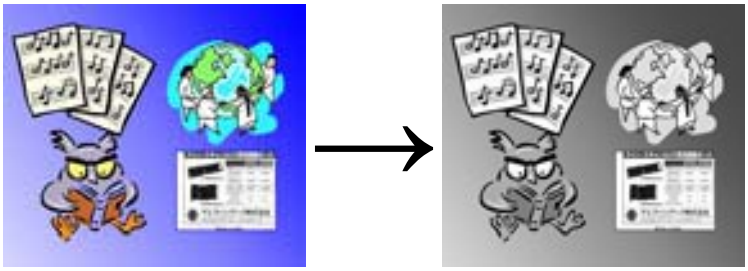
Определение зон в ручном или автоматическом режиме изолирует графические компоненты от текстовых блоков.

3. Binarization

Бинаризация преобразует серое изображение в черно-белое в соответствии с выбранными параметрами.

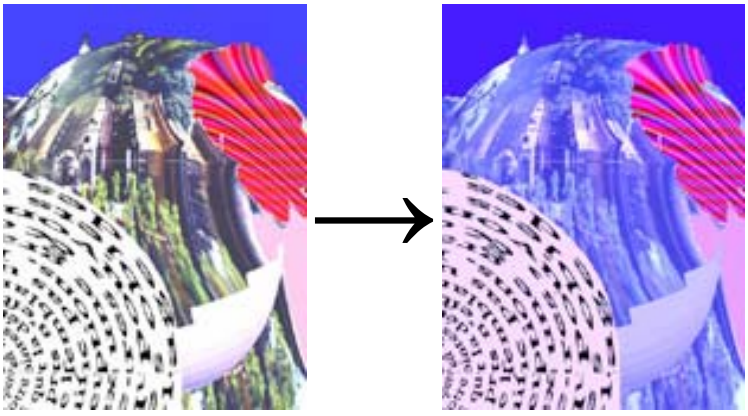
4. Color Conversion

Конвертация цветов автоматически преобразует цветное изображение в серое или черно-белое, а также наоборот.



5. Colorimetric curves

Колориметрическая кривая изменяет цветовую палитру изображения.



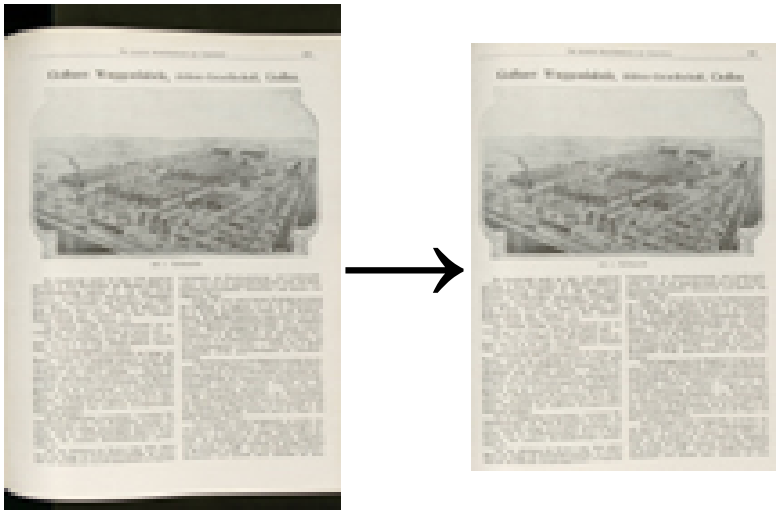
6. Crop

Обрезка позволяет:

а) разрезать изображение по вертикали или горизонтали на заданное число частей, указать положение линий разреза;

б) автоматически обрезать поля и разрезать изображения, содержащие две страницы (развороты);

в) отрезать от изображения заданные вручную поля

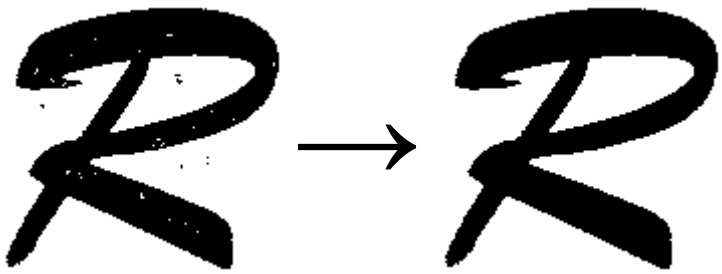


7. Deskew

Поворот изменяет угол наклона изображения, что позволяет получить правильно вертикальное расположение страницы

8. Despeckle

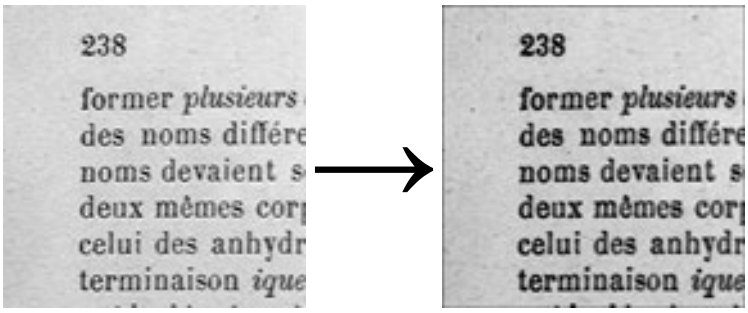
"Деспекл" удаляет отдельные белые точки ("дырки") и черные точки, а также соответствующие группы соприкасающихся пикселей ("спеклы"), появившиеся в результате неправильной бинаризации или эффектов интерференции при сканировании.



9. Filter

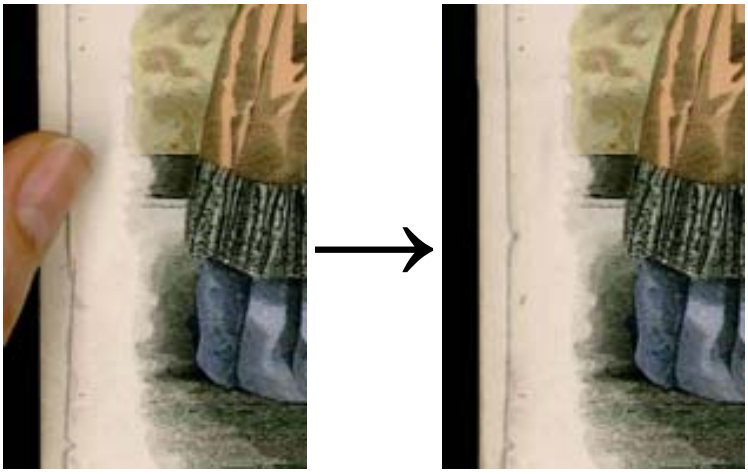
Фильтрация позволяет осуществить размытие изображения, увеличение резкости. а также утолщение или утоньшение букв. Кроме того, Book Restorer позволяет сконструировать свой собственный фильтр.

На рисунке приведен пример с утолщением.



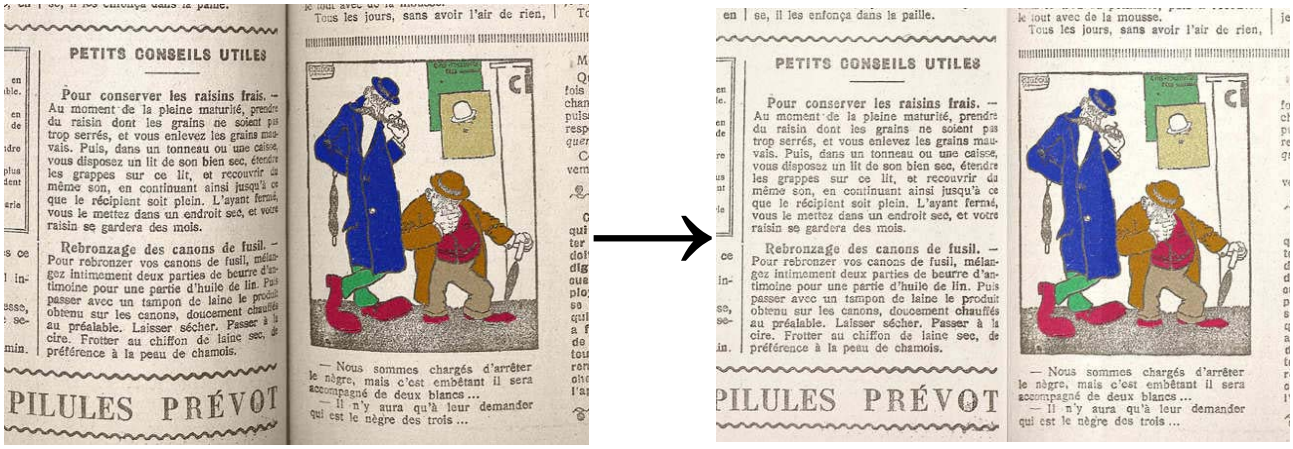
10. Finger masking

Маскирование пальца удаляет с изображения пальцы оператора планетарного сканера, держащего книгу за край. При этом маска воссоздается из аналогичных зон изображения.



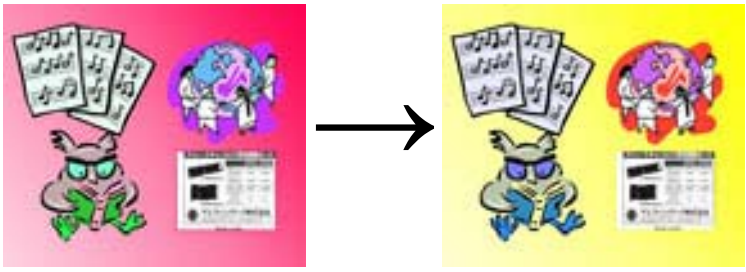
11. Geometrical correction

Геометрическая коррекция исправляет форму изображения, искривление (изгиб) листов страницы, в т.ч. у корешка книги и вследствие мятой бумаги. Принцип работы заключается в определении кривизны изгиба листа и выявления прямых элементов на странице (строчки текста, линии, поля при сканировании).



12. Hue, Saturation, Lightness

Позволяет изменять колориметрические компоненты изображения в соответствии с представлением цветов в рамках HLS-модели. Это цветовая модель, в которой цветовыми координатами являются тон (hue), насыщенность (saturation) и светлота (lightness).

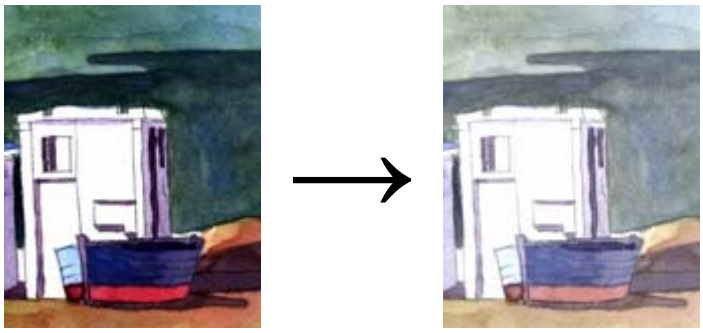


Пример: hue + 80°

13. ICC Profile

Позволяет синхронизировать оборудование, участвующее в процессе обработки изображения.

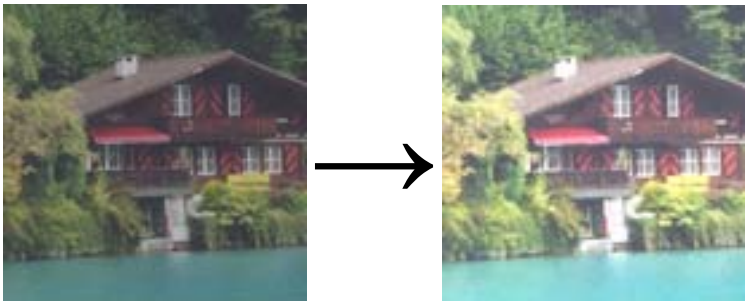
ICC профиль - это стандарт перевода информации о цвете, созданной одним устройством, в "родное" цветовое пространство другого устройства, разработанный International Color Consortium.



Для использования нужна информация о профиле устройства-источника изображения (сканер, фотоаппарат) и устройства-обработчика (монитор, принтер).

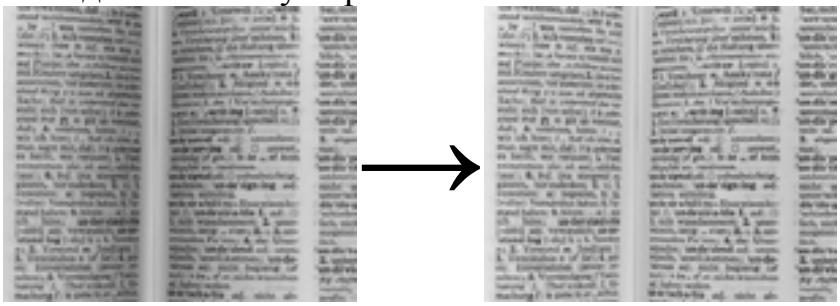
14. Light and contrast

Позволяет изменить освещенность изображения - сделать его светлее или темнее, а также скорректировать отношение яркостей наиболее темных и наиболее светлых участков изображения.

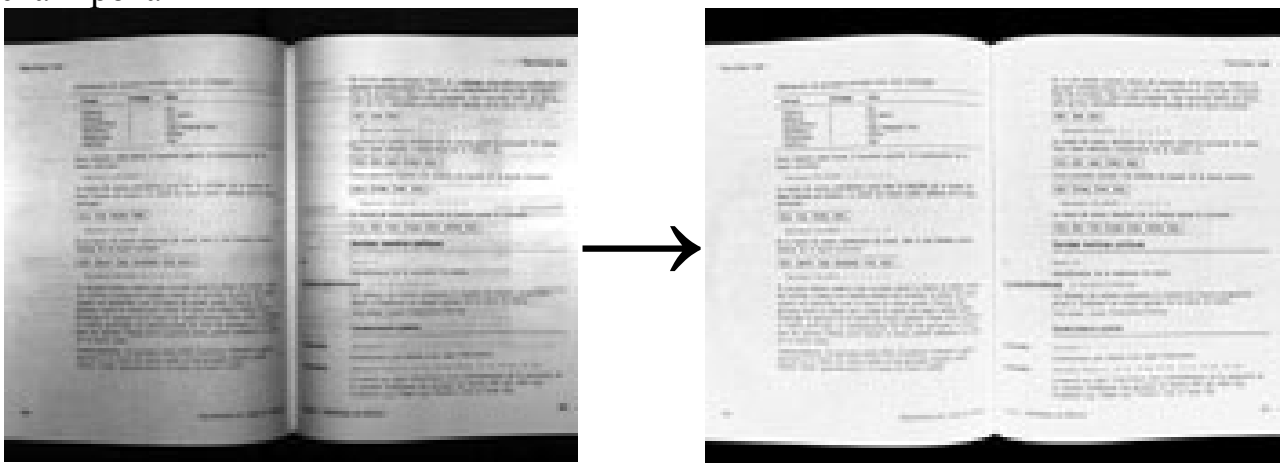


15. Lighting correction

Включает две возможности: одномерное и двумерное исправление освещенности. 1-D-коррекция направлена на исправление дефектов освещенности вследствие тени у корешка толстой книги.

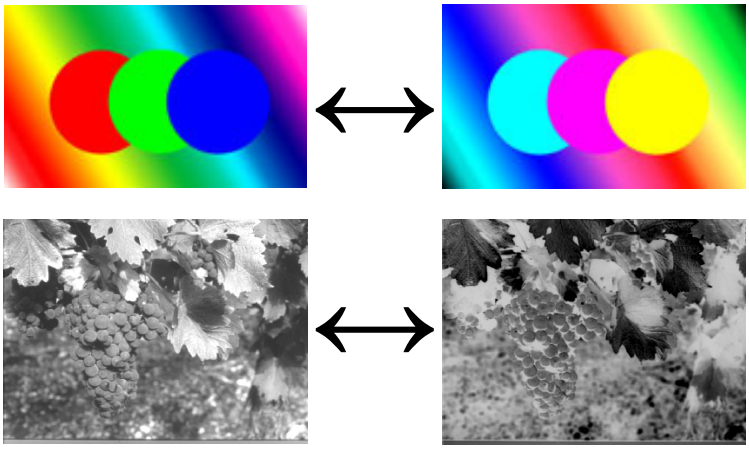


2-D коррекция направлена на исправление изменения яркости в процессе сканирования



16. Negative

Позволяет заменить цвета обратными (дополнительными).



17. Optical distortion

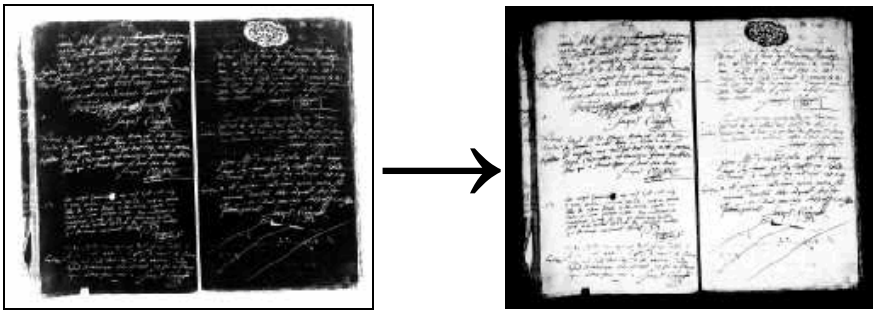
В соответствии с названием должен корректировать дисторсию (оптическое искажение), но требует наличия файла калибровки сканера Digibook и поэтому в его отсутствие неприменим.

18. Photoshop(R) script

Позволяет вызывать скрипты Фотошопа и применять их к изображениям.

19. Polarity Detection

Автоматически определяет является изображение позитивом или негативом и в последнем случае заменяет цвета обратными (дополнительными).

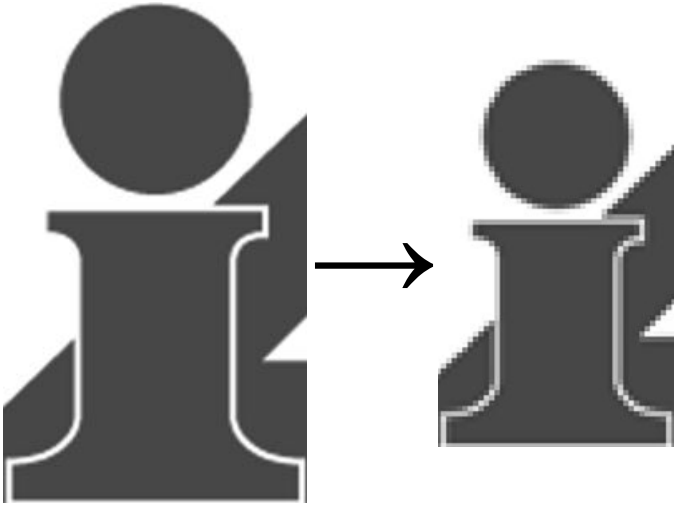


Применяется при оцифровке микрофильмов и фотопластинок

20. Resize

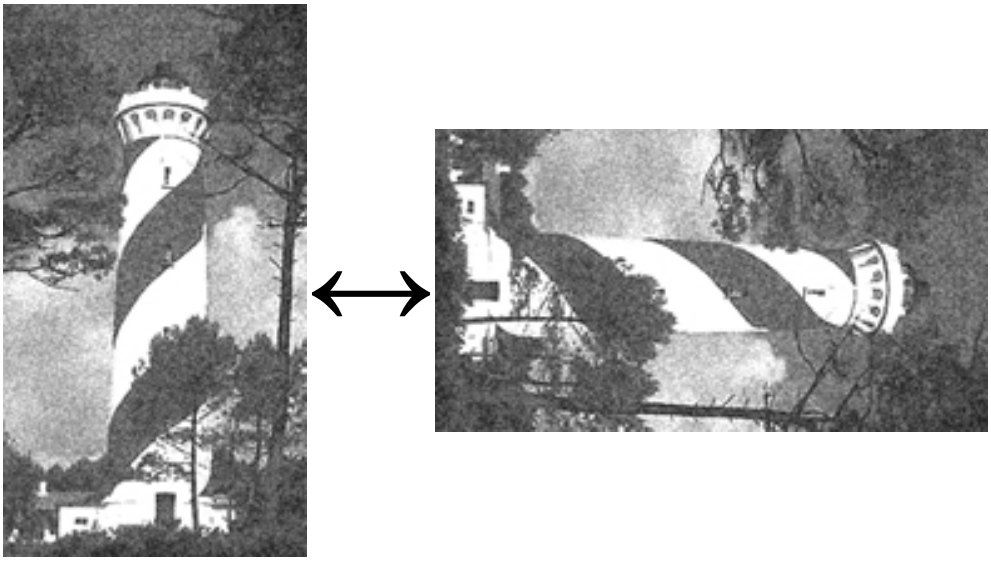
Изменение размера изображения позволяет модифицировать разрешение, размер или пропорции изображения.

Пример: снижение разрешения с 300 до 72 dpi и уменьшение размера на 40% (60% исходного).



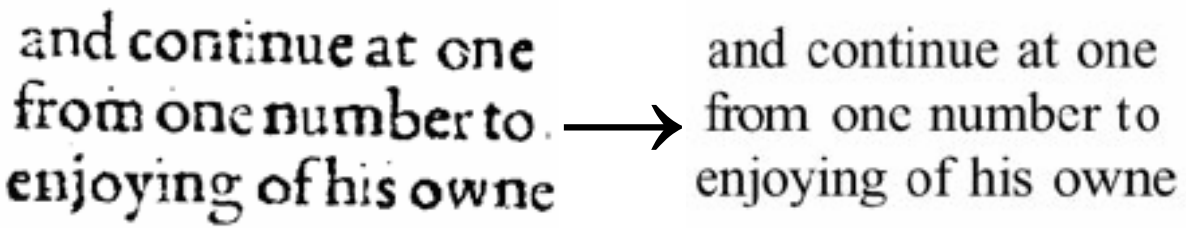
21. Transformation

Осуществляет повороты на 90, 180, -90 градусов, а также вертикальные и горизонтальные переворачивания (зеркальные отражения) изображения



22. OCR

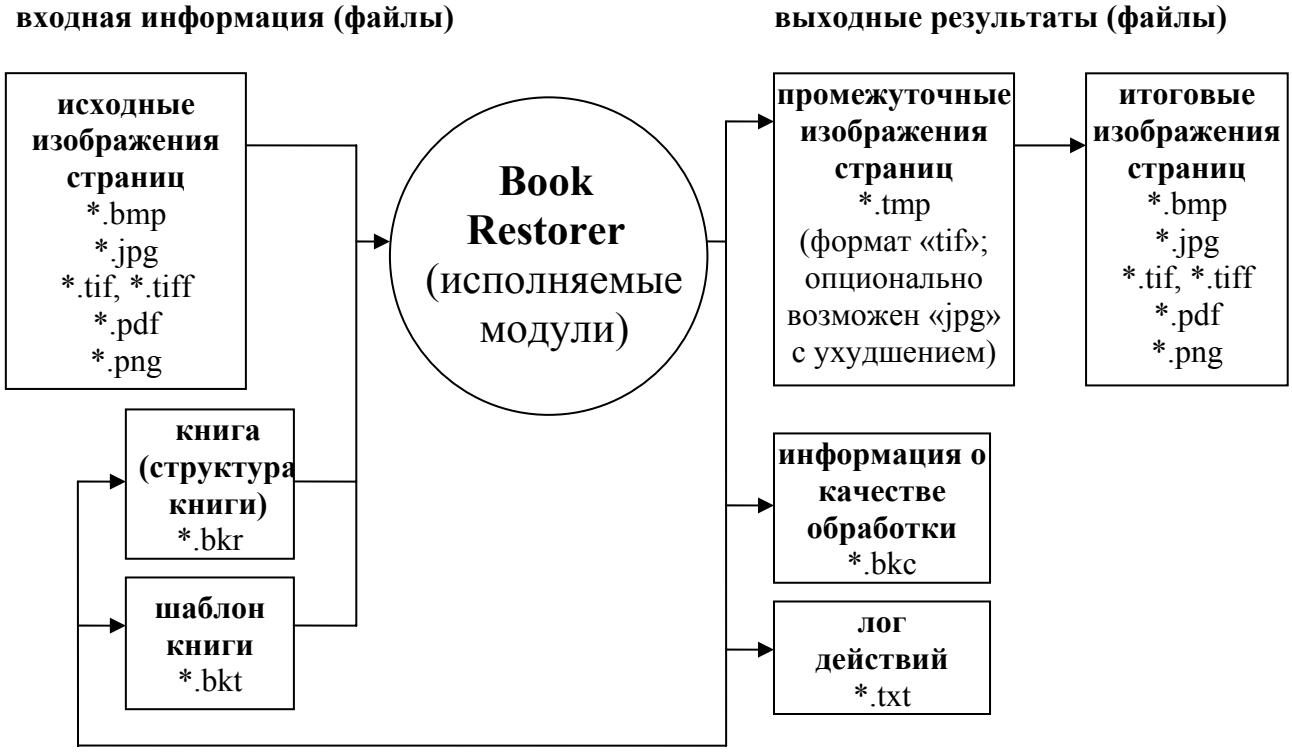
Версия 4.1. Book Restorer содержала возможность вызова сторонней программы распознавания текста, по результатам работы которой на изображении появлялся бы новый текстовый блок, содержащий распознанный текст.



1.3. Файлы и форматы Book Restorer

Программа Book Restorer предназначена для обработки отсканированных книг («реставрации»), поэтому входными и выходными файлами для нее являются графические изображения. Программа работает с файлами следующих форматов: *.tif (одностраничные и многостраничные, с разными алгоритмами сжатия G4, LZW; несжатые; 4 и 8 бит); *.jpg (а также *.jp2); *.bmp, *.png, *.pdf (в том числе векторные). Хотя это и не указано в справке, но возможно импортировать файлы Photoshop *.psd. Изображения в цветовой модели CMYK, RGB 16 bit не могут быть импортированы в Book Restorer.

В процессе работы Book Restorer создает файлы, в которых хранится информация о структуре книги, о стилях, о качестве обработки и ведется лог операций. Кроме того, Book Restorer создает «временные» файлы графических изображений, отражающих процесс редактирования сканов:



В начале работы Book Restorer для книги с указанным именем создает папку «ИМЯ_КНИГИ», в которой будет находиться три подкаталога и три служебных файла:

ИМЯ_КНИГИ
- Имя_Книги_EDITED
- Имя_Книги_ORIGINAL
- Имя_Книги_PREVIEW
  - Имя_Книги.bkr
  - Имя_Книги.bqc
  - Имя_Книги.txt

Имя_Книги_EDITED - в этом каталоге хранятся промежуточные изображения (результат применения процессов обработки);

Имя_Книги_ORIGINAL - в этот каталог могут быть скопированы или перемещены исходные изображения (Book Restorer может также оставить эту папку пустой и обращаться непосредственно к исходным изображениям, хранящимся где-либо еще);

Имя_Книги_PREVIEW - в каталоге находятся миниатюры (тумбнейлы), из которых собирается лента страниц.

В папке «ИМЯ_КНИГИ» будут также находиться служебные файлы: файл собственно книги (Имя_Книги.bkr), файл с информацией о качестве обработки (Имя_Книги.bqc) и файл лога (Имя_Книги.txt). Любопытно, что файл с отчетом о качестве обработки (*.bkc) в программе Book Restorer генерируется всегда, но просмотреть его можно только из дополнительного модуля BKR Control.

Шаблоны (файлы *.bkt) хранятся в директории, в которую проинсталлирован Book Restorer, в папке Templates. Любую нашу книгу, в которой мы определяли свои собственные стили и запоминали скрипты (наборы процессов обработки) Book Restorer может сохранить как шаблон. При этом, если сохраняется информация и о страницах, Book Restorer может создать по три указанных папки на шаблон. При создании новой книги Book Restorer предложит выбрать шаблон (из хранящихся в этой папке), из которого можно импортировать стили и скрипты.

При «публикации» книги временные файлы из папки Имя_Книги_EDITED сохраняются в выбранном пользователем формате в указанную пользователем директорию. Таким образом, помимо исходных и конечных файлов после использования Book Restorer может остаться две папки с изображениями Имя_Книги_ORIGINAL с копиями исходников и Имя_Книги_EDITED с результатами «реставрации». И если книги после обработки не удалять, то дисковое пространство расходуется со страшной скоростью. С учетом необходимости сохранения промежуточных результатов Book Restorer требуется около 4 Gb для обработки электронной книги из 200 страниц в оттенках серого при разрешении 300 dpi.

Такая работа с файлами позволяет в случае необходимости внести в исходные файлы изменения и «перегрузить» картинку, которая обрабатывается. Например, при исправлении геометрических искажений очень сильные искривления дают ошибки. Можно исходное изображение сжать/растянуть Photoshop, сбросить обработку и вернуться к «начальному» (на самом деле отфотошопленному) изображению, которое уже поддается обработке[1].

Поскольку после применения любого процесса измененное изображение записывается в директорию Имя_Книги_EDITED, то нельзя отменить отдельные шаги в обработке. Можно отказаться только от всех процессов обработки разом. Book Restorer тогда в этой временной папке заменит наше изображение на исходное.

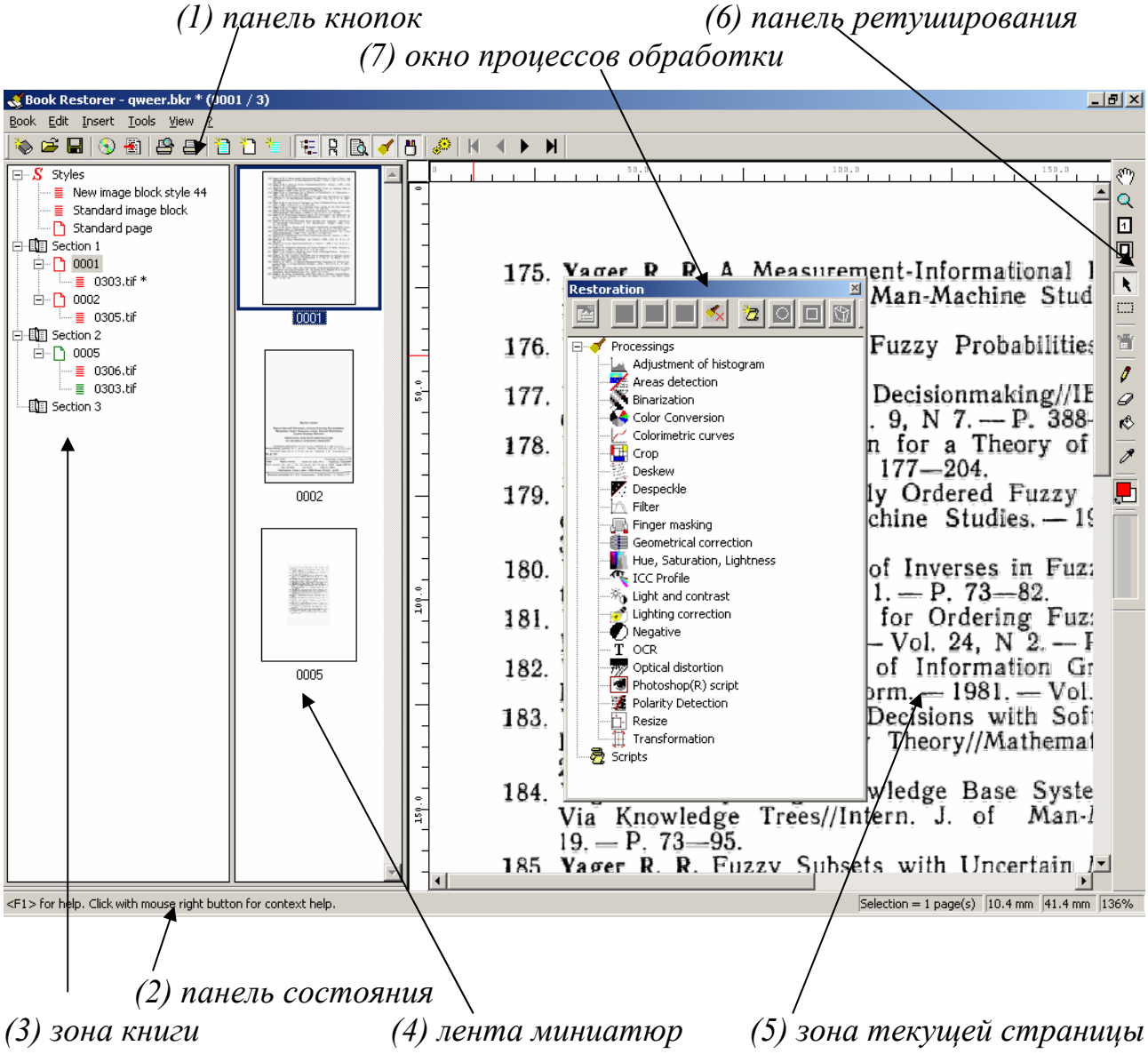
[1] После редактирования исходного файла, чтобы в папке Имя_Книги_EDITED появилось исходное изображение следует применить какой-нибудь инструмент, а потом отменить его действие - тогда Book Restorer подтянет исходник.

2. Интерфейс программы

2.1. Основные элементы интерфейса

Интерфейс БР состоит из 7 зон помимо основного меню:

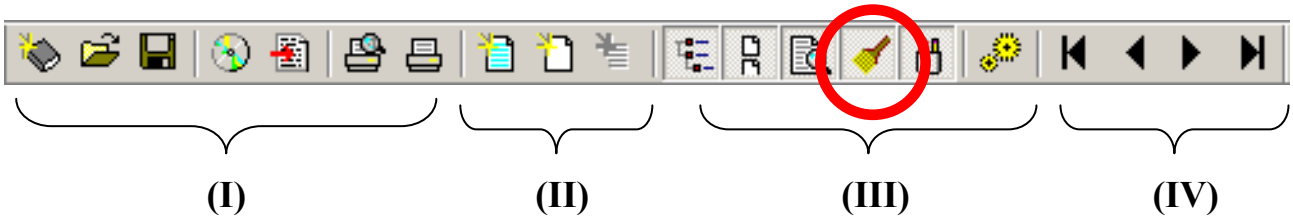
1) панели кнопок,
2) строки состояния,
3) зоны дерева книги,
4) ленты миниатюр, ,
5) зоны текущей страницы
6) линейки ретуширования
7) окна процессов реставрирования,



Самыми главными зонами являются зона книги (позволяющая выбирать обрабатываемые страницы и указывающая на их статус) и окно процессов обработки (отвечающее за выбор инструментов и их отмену).

2.2. Панель кнопок (линейка инструментов) и меню

Панель кнопок в программе Book Restorer является на редкость бесполезной. По большому счету, на ней нужна всего одна кнопка - вызов окна процессов обработки (пиктограмма в виде метелки):



Все кнопки на ней можно разделить на четыре группы в зависимости от их функций.

Во-первых, это стандартные кнопки: "создать новую книгу", "открыть существующую книгу", "сохранить книгу", "опубликовать книгу" (почему-то имеет пиктограмму в виде компакт-диска), "экспортировать лог", "предварительный просмотр печати", "печать". Но их, если мы ими пользуемся, удобнее через меню вызывать. Функция "опубликовать книгу", впрочем, является нестандартной. Под "опубликованием" в Book Restorer понимается сохранение обработанных изображений в указанную директорию.

Во-вторых, это добавление в книгу новых страниц, созданных из импортируемых изображений или пустых, а также вставка в уже существующие страницы новых блоков-изображений.

В третьих, это корреспондирующие с функциями меню кнопки сокрытия и отображения основных зон программы: зоны книги, ленты с уменьшенными изображениями страниц, изображения текущей страницы, окна с процессами обработки и панели ретуширования. здесь также можно вызвать модуль автоматизации.

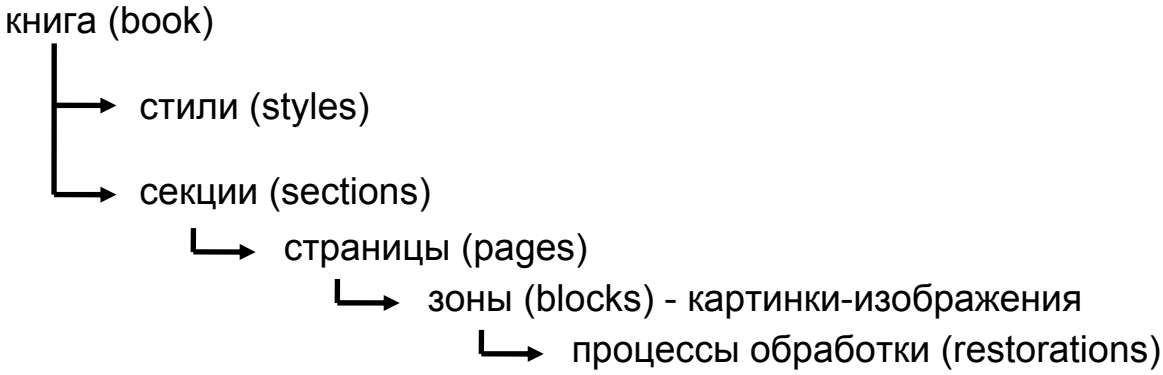
В-четвертых, это кнопки перелистывания страниц в книге. Но, на мой взгляд, перемещаться внутри книги проще и быстрее, используя дерево книги (зону книги).

Что касается меню, то в нем интерес представляют функции, которые не вошли на панель инструментов. Это, во-первых, сохранение книги как шаблона, что позволяет сохранить созданные скрипты. Во-вторых, анти-алиасинг (сглаживание черно-белых изображений), что облегчает контроль за процессом. В-третьих, наверное, это параметры настройки программы (меню Book - Preferences), которыми я, честно признаюсь, никогда не пользовался. Теоретически, в них можно настроить вызов внешней программы-редактора для изображений вместо неудобной панели ретуширования.

Рассмотрим несколько подробнее ключевые зоны - дерево книги и панель процессов обработки.

2.3. Зона книги (дерево книги)

Book Restorer имитирует структуру книги. Рабочий файл программы (*.bkr) - это "книга". В нем, и это отражает зона "дерево книги", элементы издания увязаны в иерархическое дерево со следующей структурой:



Стили - это два типа стилей: стиль страницы (размер, поля, способ нумерации) и стиль изображения (расположение на странице, размер в миллиметрах). При создании книги автоматически создаются два стиля: стиль стандартной страницы и стиль стандартного изображения. Однако, они могут быть изменены, и в дополнение к ним путем редактирования могут быть созданы новые стили пользователя.

Секция - это группа страниц. Позволяет, в частности, организовать книгу по главам или по блокам страниц с различной нумерацией. По умолчанию книга состоит из одной секции. Чтобы переместить страницу из одной секции в другую - последняя должна быть непустой. Для этого в ней может быть создана пустая белая страница.

Страница - это болванка ("канва", в терминологии Photoshop): белый лист, на котором будут расположены графические зоны - картинки-изображения. С каждой страницей ассоциирован ее стиль (размер, поля, способ нумерации). По умолчанию (в стиле "стандартная страница") размер страницы определяется равным размеру импортируемого (вставляемого) в книгу изображения.

Зона (блок) - это графическое изображение (графический файл). Для каждой зоны определены её расположение на странице и отображаемый размер (стиль изображения). По умолчанию (стиль "стандартное изображение") расположение зоны задается "по центру страницы" и она имеет "фактический" размер.

Процессы обработки - это выбранные для данной страницы (точнее для изображения) процедуры трансформации (преобразований), для которых установлены конкретные значения параметров обработки, но которые вместо немедленного применения были отложены. Для такого отложенного применения они вызываются не непосредственно, а через кнопку ("добавить в пакетную обработку") в диалоговом окне процессов обработки.

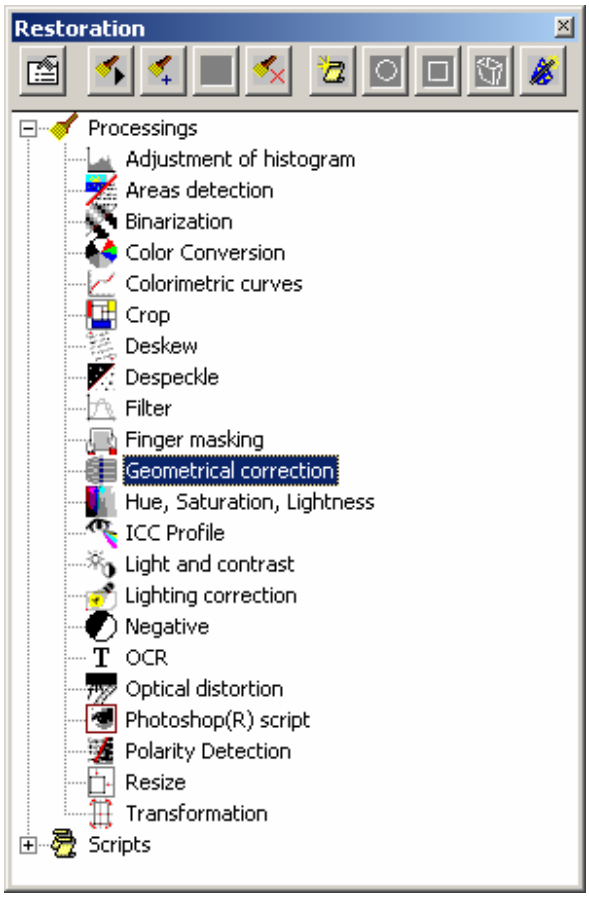
При отображении элементов книги применяются принципы цветового кодирования: зеленым цветом отображаются измененные вручную элементы, после обработки зон появляется "звездочка".

С элементами книги из контекстного меню и из основного меню Book Restorer можно делать простейшие операции. Секции можно создавать (Insert - New - Section) и удалять. При импортировании изображений можно включить опцию «создавать каждую страницу в отдельной секции».

Страницы можно удалять и перемещать (Move). Создать страницу можно путем вставки белого листа (Insert - New - White Page) или импортирования изображения страницы (Insert - Import files). Страница может одновременно содержать блоки (зоны) различного формата (tif и jpg, разное число цветов). При наличии нескольких блоков на странице их можно выделить в отдельные страницы (Edit - Divide in several pages).

Блоки (зоны) можно создать на уже существующем изображении страницы (инструмент Areas detection), можно вставить новые изображения в выбранную страницу (Insert - Add block). Можно задать расположение блоков относительно друг-друга при их наложении (Edit - Order), при этом изменяется порядок блоков в дереве . Блоки можно объединять (Edit - Flatten blocks).

2.4. Окно процессов обработки ("реставрирование")



Окно процессов обработки (панель «реставрирование») появляется в рабочем окне сразу после загрузки приложения Book Restorer.

Если это окно закрыто, снова его открыть можно нажатием на кнопку панели инструментов. Запуск процедур обработки может быть осуществлен также через меню «Tools - Restore».

В окне содержится древовидный список процессов обработки, в нижней части которого видны «скрипты» - формируемые пользователем наборы процессов (с тем чтобы их можно было, однажды подобрав параметры, применять к разным страницам). Как процессы так и скрипты можно сразу применить к выбранным страницам, а можно, закрепив процесс за выбранной страницей, обработку отложить.

Данное окно также содержит панель кнопок, которые являются активными (неактивными) в зависимости от выбранного объекта. Причем от объекта, выбранного не только в «дереве процессов», но и в «дереве книги». Если бы все кнопки были активными, то панель выглядела бы следующим образом:

Находящиеся на данной панели кнопки выполняют следующие функции:

- «Settings» - для выбранного процесса открывает окно настройки параметров. (Используется при редактировании процессов внутри скриптов, т.к. для отдельного процесса нажатие кнопки ОК в окне выбора параметров в этом случае не приводит к его запуску.) Если процесс не требует параметров (например, Negative), то нажатие на кнопку игнорируется. Для выбранного скрипта открывает окно, в котором редактируется название (имя) скрипта;

- «Apply Now» - запускает для страницы применение выбранного процесса обработки или скрипта. Если процесс требует параметров, то открывает окно их выбора, и запуск процесса осуществляется при нажатии кнопки ОК в этом окне. Применяется к выбранным страницам в дереве книги;

- «Add for batch processing» - добавит для выбранной страницы выбранный процесс в отложенную обработку (появится в «дереве книги»);

- «Do batch processing» - применит для выбранной страницы все отложенные процессы. (Процессы исчезнут из «дерева книги»);

- «Original» - возврат к исходному изображению и отмена всех примененных процессов. Обратите внимание, что отменяются все процессы, а не только последний, что составляет проблему запоминания их параметров;

- «New Script» - создаст новый скрипт с именем «New Script #», не содержащий каких-либо процессов обработки;

- «Start script recording» - для выбранного скрипта при данной нажатой кнопке все вызываемые процессы будут добавляться в скрипт. Записываемый скрипт при этом помечается красным кружочком;

- «Stop script recording» - останавливает запись скрипта;

- «Delete» - позволяет удалить скрипт полностью или любой из процессов внутри скрипта (например, с целью изменения их очередности);

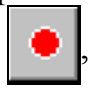
- «Script wizard» - вызывает странный «мастер скриптов», который позволяет выбрать один из четырех предустановленных скриптов, предназначенных для архивации или печати черно-белых или цветных страниц:

название скрипта	«To archive in binary»	«To archive in grey levels/color»	[To print] «Originals in binary»	[To print] «Originals in grey /colours»
процедуры в составе скрипта	Lighting correction Filter Binarization Despeckle Crop Resize	Lighting correction Geometrical correction Crop	Despeckle Redressement (Deskew) Crop	Lighting correction Geometrical correction Binarization Despeckle Crop

Таким образом, данное окно процессов обработки (панель «реставрирование») позволяет осуществлять три основных задачи:

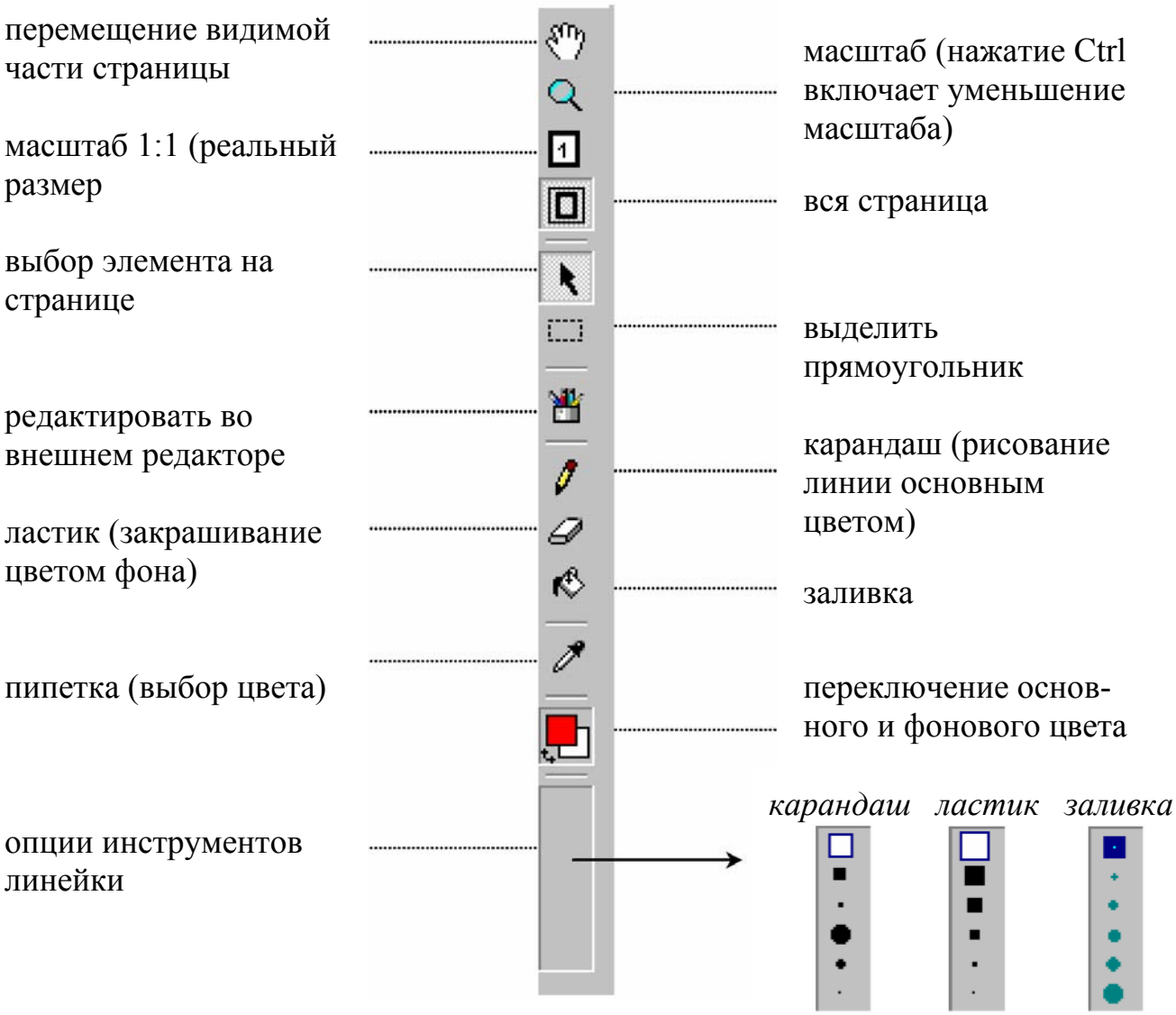
1) Немедленное применение процесса обработки к изображению. (Двойной клик на процессе, или нажатие на кнопку , выбор параметров процесса обработки в диалоговом окне, нажатие в диалоговом окне кнопки ОК).

2) Отложенное применение процесса обработки к изображению. (Нажатие на кнопку , выбор параметров процесса обработки в диалоговом окне, нажатие в диалоговом окне ОК. Применение процесса для выбранных страниц осуществится при нажатии на кнопку или из модуля BKR Automation.)

3) Создание скрипта - набора последовательных процессов обработки (Нажатие на кнопку , затем на  , выбор процесса, двойной клик на нем, выбор параметров процесса и нажатие кнопки ОК в окне, далее кнопка .)

2.5. Панель (линейка) ретуширования

Данная панель в меню называется "retouch toolbar", на панели кнопок "touch-up toolbar". В файле справки ее также называют "retouch palette". Я такое редактирование уже привык в Scan Kromsator делать. Наверное, ластик здесь можно использовать:



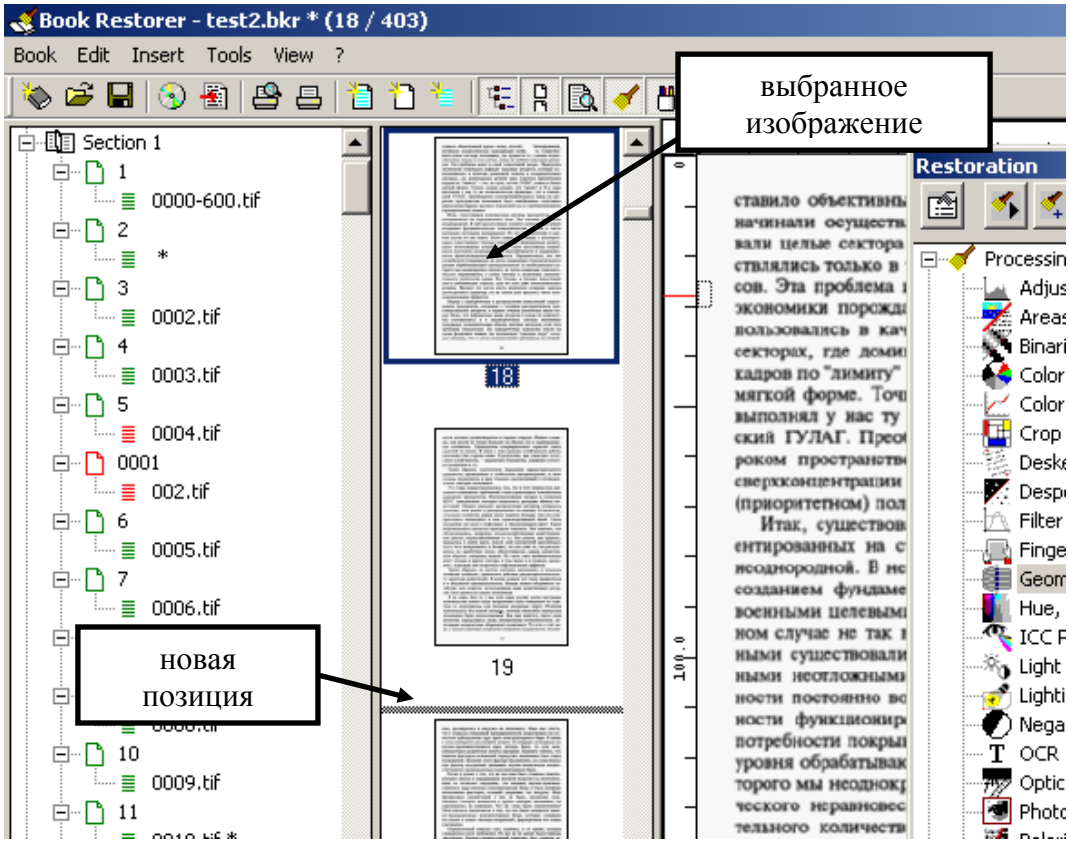
Панель Retouch содержит «стандартный» набор средств графического редактирования для рисования, стирания, изменения цвета отдельных пикселей изображения. Позволяет копировать цвет отдельных пикселей. Размеры, форма и чувствительность таких средств, как **карандаш**, **ластик** и **заливка** могут быть изменены. Из нее также может быть вызвана задаваемая в настройках Book Restorer внешняя программа редактирования изображений.

Команда *Edit - Undo retouch* (или с клавиатуры Ctrl+Z) отменяет выполненные изменения. «Запоминает» при этом десять последних операций.

Панель также содержит инструменты манипулирования изображением страницы: перетаскивание, увеличение/уменьшение (переключение происходит при нажатой клавише "Control"); автоматическая подгонка изображения под размеры поля; отображение в оригинальном размере.

2.6. Лена прокрутки страниц, зона страницы и строка состояния

Важной особенностью Book Restorer является наличие такого элемента как лента прокрутки страниц, в которой отображаются текущие (с учетом применения инструментов обработки /"реставрирования"/) изображения страниц. Это значительно облегчает контроль за подготовкой книги, так как позволяет сравнивать страницы друг с другом. Однако, выключение ленты значительно ускоряет работу программы. Помимо осуществления визуального контроля лента позволяет легко и просто перемещать страницы внутри книги путем перетаскивания. Для этого необходимо кликнуть по нужной странице и, не отпуская кнопку мыши, переместить ее к целевому положению, в котором появится серая линия. После чего отпустить кнопку мыши.



Справа от ленты (в зоне страницы) укрупнённо отображается текущая страница, к которой можно применить инструменты "ретуширования" (обычного редактирования изображения). Отображаемая страница может быть увеличена или уменьшена с панели ретуширования. Для неё из меню также может быть включен анти-алиасинг - сглаживание черно-белых изображений.

Внизу рабочего окна Book Restorer расположена строка (панель) состояния. В ее левой части показывается абсолютно бесполезная информация о выбранной функции или разделе меню. А вот в ее правой части:

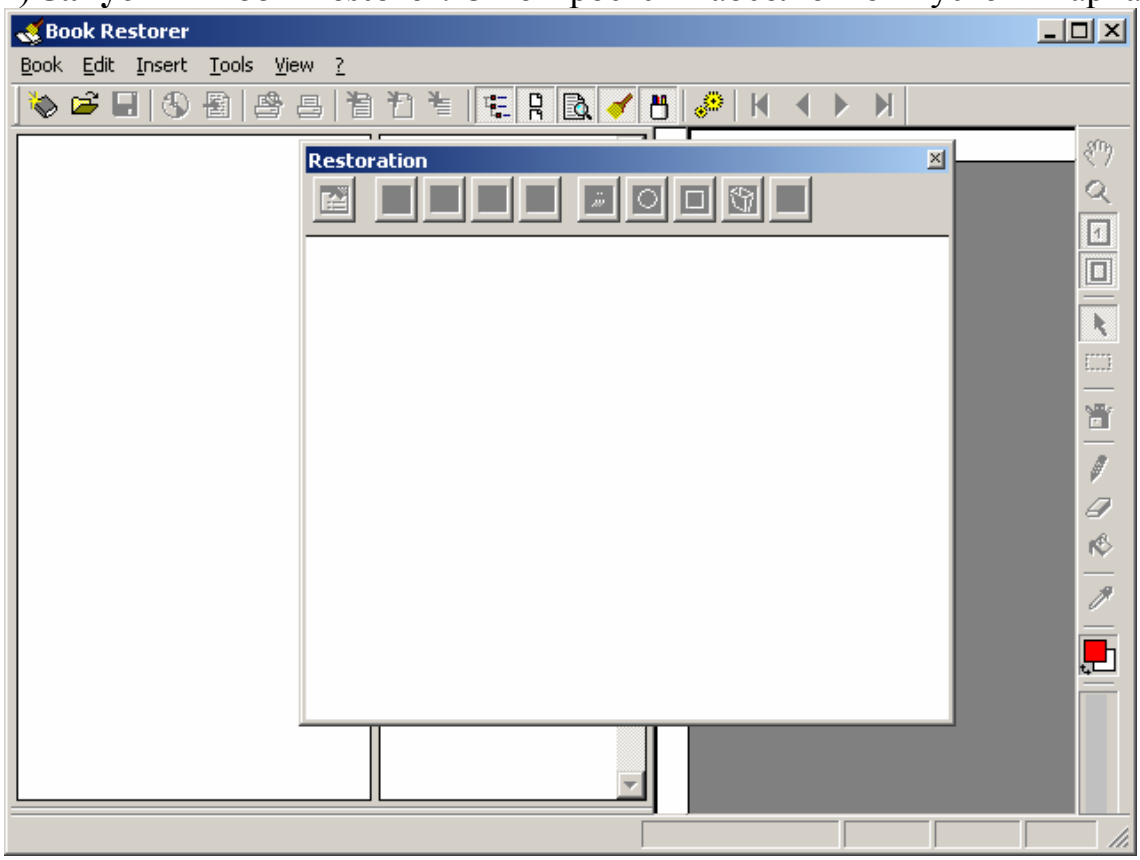
дается информация о количестве выделенных страниц или блоков, о положении курсора (когда он находится на текущей странице) относительно верхнего левого угла страницы и о масштабе отображения текущей страницы

3. Использование Book Restorer

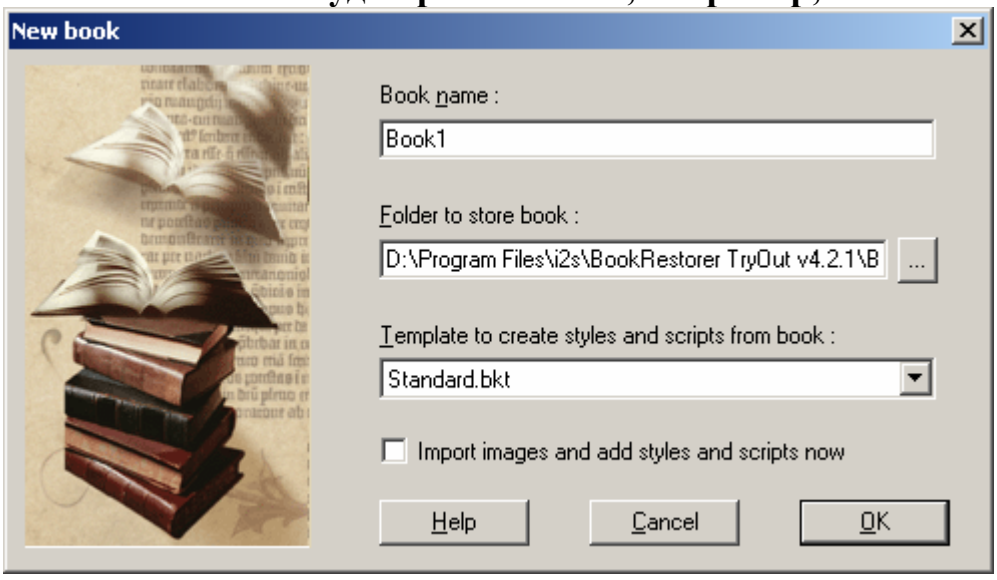
3.1. Создание книги и «опубликование» обработанных страниц

Рассмотрим процесс обработки «сырых сканов», полученных на обычном сканере и сложенных в папке «Scans\Book1».

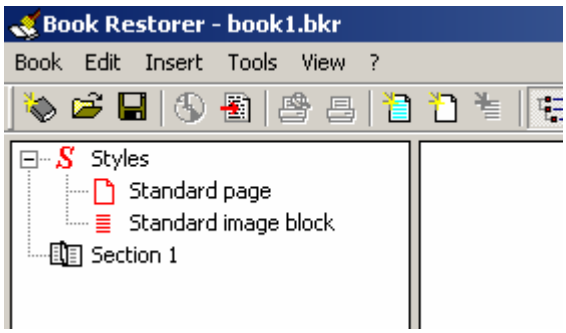
1) **Запустим Book Restorer**. Он откроется в абсолютно «пустом» варианте.



2) **Выберем команду «создание книги» (меню Book → New), и зададим нашей книге какое-нибудь красивое имя, например, «Book1»**:



Укажем его в первом поле открывшегося диалогового она и нажмем кнопку «ОК» - создастся книга с таким именем. В окне программы при этом появится пустое дерево книги:



На жестком диске при этом появится каталог BOOK1, внутри которого будет расположен собственно файл книги book1.bkr. Этот каталог создастся внутри папки, которая отображена во второй строке диалогового окна. По умолчанию там указана директория, в которую установлен Book Restorer. Если эта директория не подходит, то можно задать другую, кликая на кнопку с многоточием. В папке BOOK1 также создадутся подпапки для хранения промежуточных изображений и некоторой другой информации (см. § 1.3.)

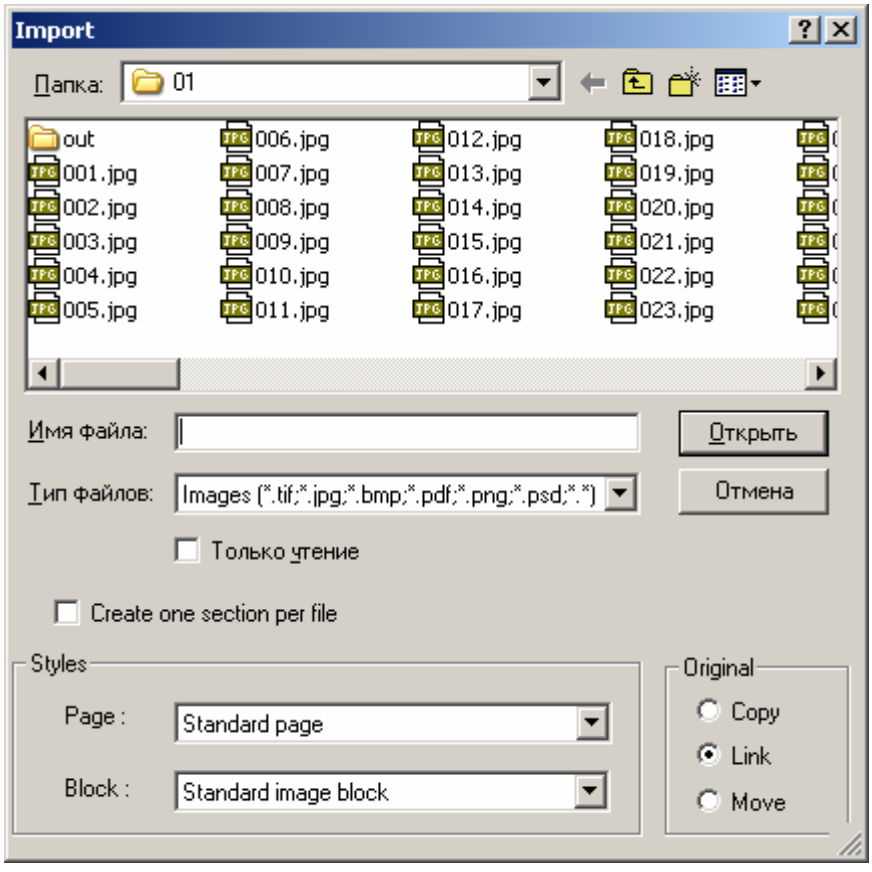
В этом же окне предлагается выбрать шаблон (по умолчанию задан «стандартный). Можно также выбрать «отсутствие шаблона» или выбрать свой шаблон, из которого будут импортированы стили (если есть заранее созданный). Нам это при обычной обработке делать смысла не имеет.

Кроме того, в этом диалоговом окне есть галочка «*Import images and add styles and scripts now*». Если её нажать, то запустится мастер (точнее помощник) по составлению книги: а) сначала откроется окно по выбору или редактированию стилей; б) затем окно по импорту изображений; в) затем окно по выбору скриптов. Нам это тоже всё не нужно.

3) **Импортируем изображения страниц в нашу «книгу»** (то есть в «пакет», объединяющий структуру книги, стили, нумерацию страниц, изображения, процессы обработки и т.д.).

Для этого в меню выберем команду: ***Insert → Import files***, и выберем нужные файлы.

Нажмем кнопку «Открыть», и в дереве книги (слева в основном окне программы мы увидим список страниц и связанные с ними изображения.



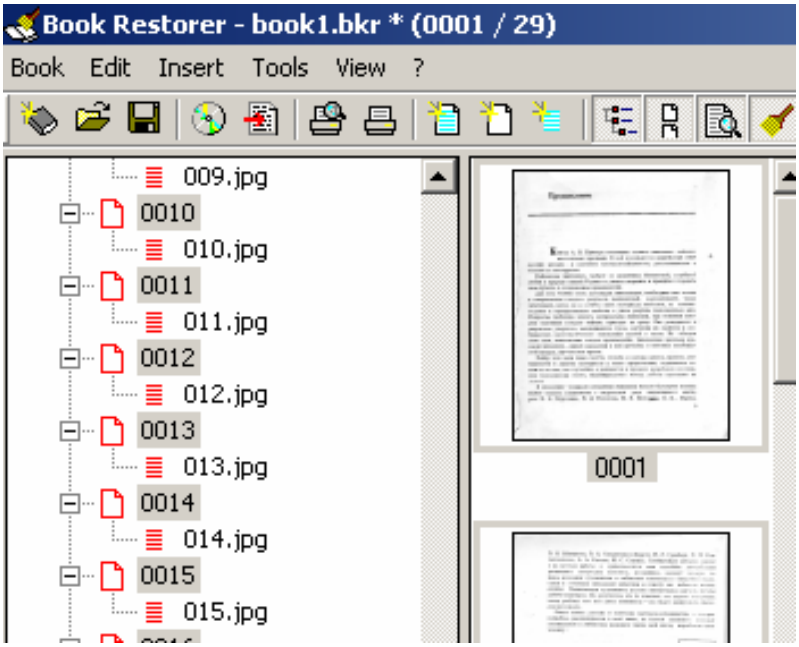
Поскольку мы не изменяли предлагаемый по умолчанию и для страниц и для импортируемых основных изображений «стандартный» стиль, то все наши элементы книги обозначаются красным цветом.

В диалоговом окне, которое открывается при импорте файлов, существует возможность выбора других стилей для страницы и импортируемого изображения. Интереса для нас, если мы не занимаемся потоковым сканированием, это тоже не представляет. В этом диалоговом окне также существует возможность ограничить выбор только каким-то определенным типом файлов, существует опция «создавать каждую страницу как отдельную секцию» (галочка);

Что может быть интересным, то это радиокнопка, отвечающая за то, что будет происходить с исходными изображениями. По умолчанию стоит опция «Link». То есть в нашей книге будет сохранена информация, где находятся сканы. Но если наши сканы вдруг будут удалены или перемещены, а нам потребуется изменить нашу книгу, это приведет к проблеме. (Я лично часто сталкиваюсь с такой ситуацией: где исходники - неизвестно.) Можно выбрать опцию «Copy» тогда исходники будут скопированы в ту папку, где находится наша книга и промежуточные результаты обработки (но это приведет к существенному увеличению занимаемого дискового пространства). Можно выбрать «Move», тогда они будут перемещены, и есть риск не найти их на прежнем месте, если папки, куда осуществляется сканирование, имеют постоянное размещение, а папки Book Restorer то создаются, то удаляются.

В Book Restorer также существует возможность задать папку, из которой следует импортировать все изображения. При этом к ним даже можно применить какой-либо заранее сохраненный в шаблоне скрипт. Это делается командой меню *Insert - Automated Import:*

4) **Выделяем все изображения в нашей книге (те изображения, которые хотим обработать).** Из меню командой *Edit - Select All:*



Страницы будут выделены серым или синим цветом. Если мы хотим обработать не все изображения, то выделяем в дереве книги только необходимые. При нажатой на клавиатуре клавише Ctrl можно выделять и снимать выделение отдельных страниц.

Если же при многих выделенных страницах кликнуть на одной без зажатой клавиши Ctrl, то выделение остальных снимется.

5) Применяем необходимую обработку.

Выбираем необходимый нам процесс из меню: *Tools - Restore - «процесс»* или делаем на нём двойной клик в окне процессов обработки (реставрации). Если окно процессов обработки не открыто, то его можно вызвать, кликая по кнопке с метёлкой на панели инструментов или выбирая из меню команду *View - Workspace - Restore Toolbar*. Это откроет окно «Restoration».

Задаем если необходимо параметры процесса, нажимаем кнопку «ОК» в его диалоговом окне, и процесс будет применён ко всем выделенным изображениям. (Некоторые процессы применяются сразу при их выборе, например, «Negative».) У обработанных изображений в дереве книги после их имен появятся звездочки.

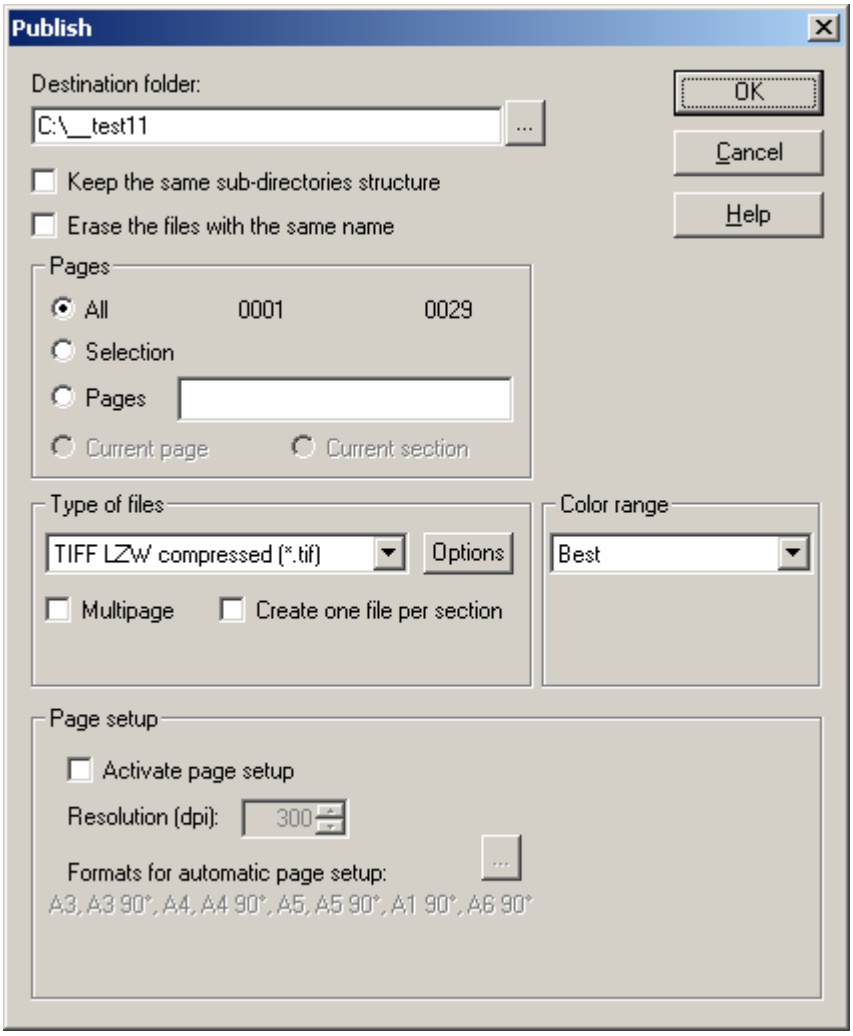
Если отключить ленту, то программа не будет пытаться постоянно перерисовывать миниатюры, и обработка пойдёт быстрее. Перемещаться потом между изображениями для целей контроля корректности выполненной обработки можно будет с помощью дерева книги, а обработанные изображения при этом будут демонстрироваться в главном окне Book Restorer справа (в т.н. окне страницы). Для неудачно получившихся изображений отменяем обработку (кнопка в окне процессов реставрации или из меню *Tools - Reload original*) и пытаемся выполнить наш процесс с другими параметрами.

6) Изменение нумерации страниц в случае необходимости.

Если вы впервые используете Book Restorer и следуете предшествующим указаниям, то скорее всего ваши стандартные страницы пронумерованы 1, 2, ... 10, ... 99, 100 (такова нумерация по умолчанию в Book Restorer после его инсталляции). Эта нумерация страниц в последующем воспроизведется при сохранении результатов (итоговых изображений после обработки) и может привести к ошибкам в очередности страниц при их обработке в других программах. Для того чтобы страницы нумеровались 0001, 0002, 0003, и т.д., следует щелкнуть правой кнопкой мыши на стиле «стандартная страница» («Standard page» в разделе «стили» в дереве книги), выбрать закладку нумерация (Numbering), и выбрать нужный формат. После нажатия на кнопку «ОК» это свойство стандартного стиля не только применится ко всем «стандартным» страницам, но запомнится и будет применяться в последующем. Так что этот пункт требует только однократного выполнения.

7) «Опубликование» результата: сохраняем итоговые результаты (обработанные изображения) в нужном нам каталоге в нужном формате. Это не печать (Print), когда изображения уходят на принтер и не сохранение (Save), когда сохраняется наша «книга» - пакет с информацией об исходных изображениях, их свойствах, и процессах обработки. Это по аналогии с типографским книгопечатанием именно «опубликование». При этом конечные результаты сохраняются не в стандартной директории внутри папки книги (как это сделано в Scan Kromsator и Scan Tailor), а в самостоятельной папке.

Для вызова диалогового окна «Publish» можно кликнуть по кнопке с компакт диском на панели инструментов или выбрать из меню команду *Book - Publish*:



Нам в этом окне следует задать целевую папку (Destination folder), оставить радиокнопку в позиции «All» (все файлы) и нажать кнопку «OK».

По умолчанию в указанной нами папке будут созданы файлы *.tif с внутренним сжатием по алгоритму Лемпеля-Зива.

Это оптимально для дальнейшей работы и создания файлов в форматах *.pdf, *.djvu.

Единственное, что хотелось бы оговорить особо - это значение в поле «Color range». Оно задает цветность. По умолчанию стоит «Best».

Это означает, что Book Restorer выберет оптимальную цветность для каждой страницы. У меня, когда я только начинал использовать Book Restorer, часто получалось так, что черно-белые страницы в итоге оказывались серыми. Я сейчас не знаю, что я тогда неправильно делал, но для контроля итоговой цветности и её исправления мне часто приходилось привлекать дополнительные программы. Мне хотелось бы посоветовать либо задавать здесь цветность в явном виде, либо не лениться лишний раз перепроверить результат.

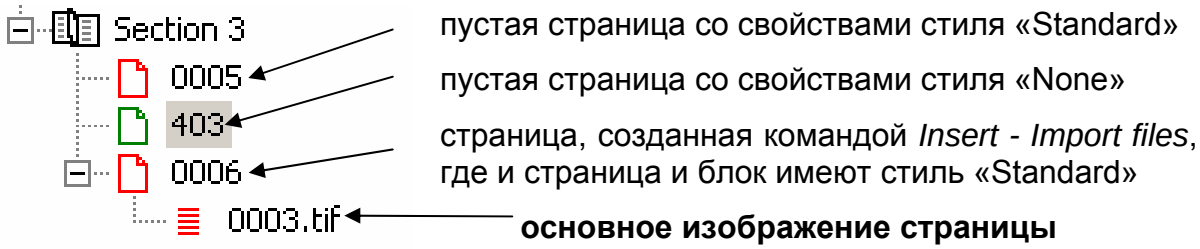
В случае, если вы всё «опубликовали», проверили и решили ряд страниц переделать (и печатаете уже не «все» страницы, а отдельные и заново), то следует поставить галочку в поле «Erase the files with the same name» - удалить файлы с одинаковыми именами. Если ее не поставить, то в дополнение к файлу 00##.tif в целевой директории будет создан файл 00##_1.tif, при повторной переделке - файл 00##_2.tif и так далее. «Публиковать» можно и несколько страниц одновременно, задавая в поле «Pages» выражения типа: «1; 3; 5-9; 14».

Данное диалоговое окно позволяет для целевых файлов задавать другие способы сжатия формата tif, выбирать конечный формат *.jpg (с указанием уровня «качества »), а также другие форматы (в том числе *.pdf).

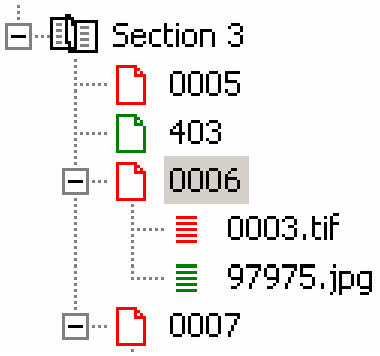
3.2. Определение и редактирование зон в Book Restorer

При начале работы в программе Book Restorer создается книга, в которой основным элементом являются страницы. Страница, как я уже неоднократно упоминал, это изображение со свойствами - поля, нумерация, и т.д. Страница может быть пустой (белой) или быть создана из уже имеющегося изображения. То есть, фактически, страница создается из «основного» изображения - пустого (белого) или полученного в результате сканирования.

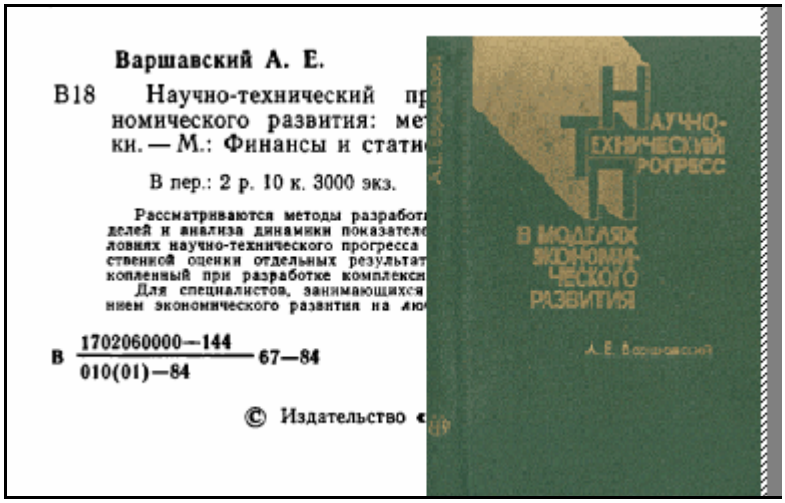
В «дереве книги» это основное изображение будет идти первым после имени страницы:



Структурно в «страницу» помимо назначенных ей свойств («стиля») входят так называемые блоки (blocks), которые представляют из себя отдельные изображения. Как в пустую страницу, так и в страницу, в которой уже существует основное изображение, можно добавить новые блоки-изображения из уже существующих файлов с помощью команды *Insert - Add a block.* (При этом в дереве книги должна быть выделена страница, к которой идет добавление, а не какой-либо из блоков).



на изображение страницы книги (0003.tif) наложено изображение обложки книги (97975.jpg). Оно выровнено «справа и вниз», поэтому стиль блока уже не «Standard»

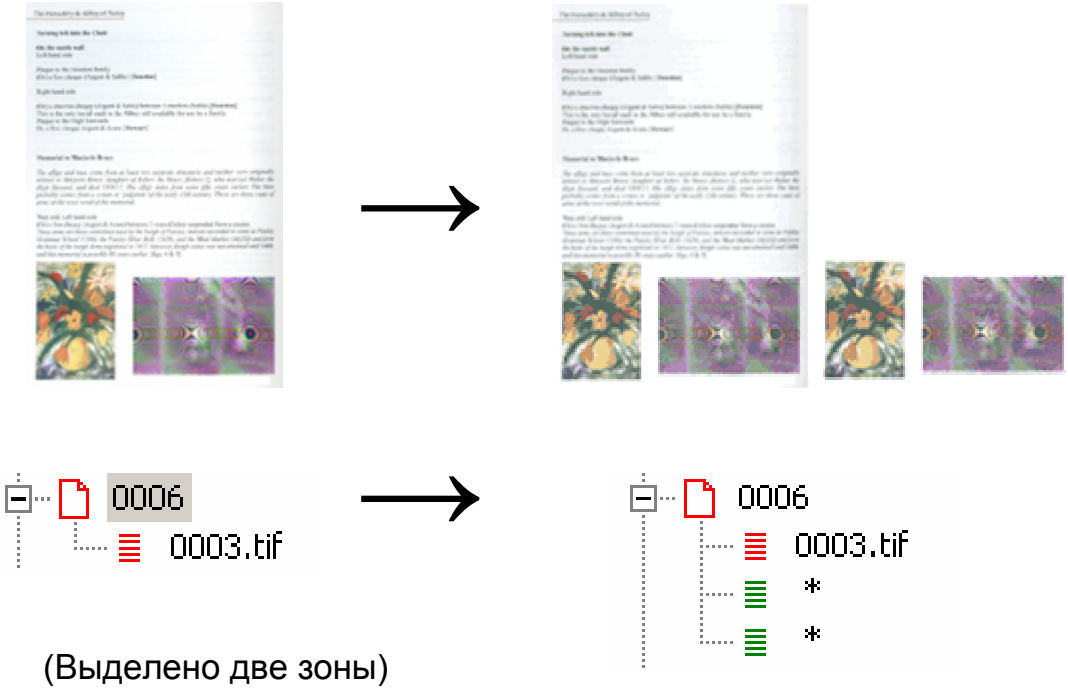


Результирующее изображение будет формироваться послойно: сначала самый первый блок, потом на него будут наложены идущие за ним в «дереве книги». Последовательность блоков в дереве можно менять командой *Edit - Order*.

Для пустой (белой) страницы можно задать фиксированный размер (изменив стиль), а затем в этот шаблон вставлять отсканированные изображения. В том случае, если все сканы страниц имеют разный размер, это позволяет сделать их одинаковыми. (Не обрезать под единый размер, а

«нарастить» до него. При этом каждому из них позже вручную можно будет задать уникальное смещение, если это, например, страницы с разными полями).

Однако помимо вставки блоков, являющихся самостоятельными изображениями, в страницу также можно вставить блоки, являющиеся зонами (участками) уже существующего изображения. Для этого используется процесс обработки «Areas detection»:



(Выделено две зоны)

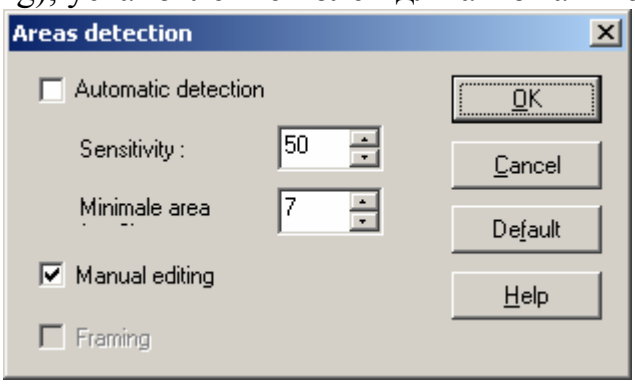
При этом на жестком диске для каждой такой найденной зоны в папке Имя_Книги_EDITED (в которой хранятся промежуточные результаты) будет создан самостоятельный файл. Это позволяет, кстати, редактировать эти промежуточные изображения с помощью внешних редакторов. Никаких других изменений с уже существующими файлами в данном «процессе обработки» не происходит. Однако дальше к ним могут применяться разные процессы.

Такое выделение зон поэтому достаточно часто оказывается очень удобным при обработке книги, потому что различные участки страницы нередко должны обрабатываться по разному. Например, текст и иллюстрации. Это также может касаться участков с различной освещенностью, участков с различной цветностью, с различными нелинейными искажениями.

Процесс обработки «Areas detection» может определить зоны автоматически (каждый элемент с высоким контрастом с фоном и размер которого больше, чем у букв, будет определен как графический). Каждая графическая зона будет связана с прямоугольным блоком, который будет скреплен со страницей. Зоны также могут быть заданы вручную. Можно попробовать определить зоны автоматически, а потом лишние вручную удалить, а недостающие - добавить.

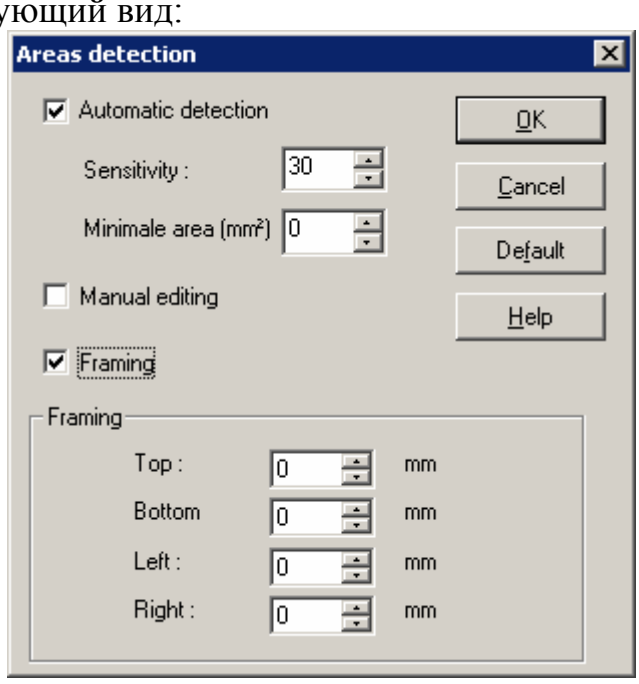
Я, впрочем, однажды получил по сорок зон на страницу и с тех пор стараюсь больше с автоматическим определением не связываться.

При выборе данного процесса обработки появляется окно, содержащее три чекбокса: осуществление автоматического определения зон (automatic detection); открытие диалогового окна для определения зон вручную (manual editing); установление полей для автоматически определенных зон (framing):



В этом окне также возможно задание параметров процесса автоматического определения зон: чувствительности и минимального размера зоны (выражен в квадратных миллиметрах). «Чувствительность» изменяется от 1 до 100 и определяет пороговое значение в черно-белом изображении по которому определяются зоны, чем этот параметр больше - тем больше зон будет определено. Это достаточно важные параметры, потому что получение неразумно большого количества зон, которые на самом деле отсутствуют и подлежат ручному удалению, неэффективно.

Последняя опция (framing) доступна только, когда выбрано автоматическое определение зон, и фактически тоже является параметром автоматического определения. Если эти опции выбраны, то наше диалоговое окно принимает следующий вид:



Зона, определенная без фрейминга

МОСКВА «НАУКА» 2000

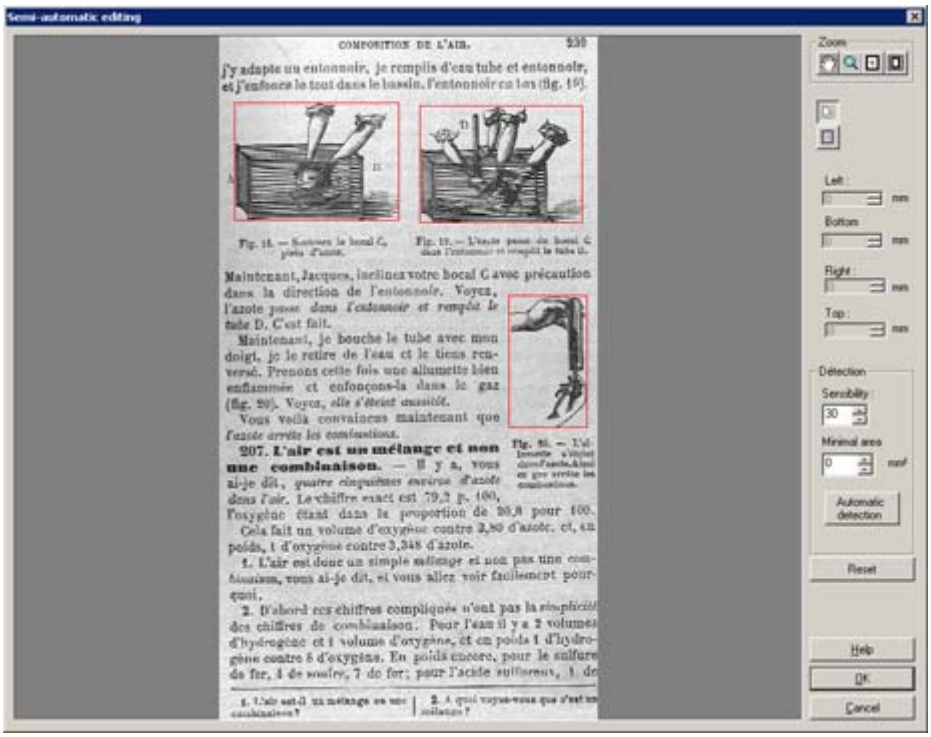
Зона определенная с фреймингом: поля по 10 мм

МОСКВА «НАУКА» 2000

В нем становится возможным задание полей на границах прямоугольников детектированных зон. То есть вокруг каждой такой зоны как бы задается рамка, и всё, что внутри нее находится, будет включено в зону. Единая для всех зон ширина этой рамки, выраженная в мм, и задается в этом окне.

Если выбрано только автоматическое определение зон (с фреймингом или без фрейминга), то после нажатия в диалоговом окне кнопки «ОК» в дереве книги появятся значки определенных (детектированных) зон. Редактировать эти зоны (уточнять границы и положение) уже невозможно - только удалять.

Если выбрано только определение зон вручную, то откроется диалоговое окно редактирования зон. Если одновременно выбраны опции автоматического определения зон и ручного редактирования, то детектированные зоны будут отражены в этом окне:



В данном окне, впрочем, доступен запуск автоматического определения зон, и если они не были определены ранее или вы вдруг удалили эти зоны в процессе редактирования, они могут быть детектированы заново.

Не активируйте, впрочем, ручное выделение зон при потоковой обработке, иначе на каждом изображении страницы процесс будет прерываться и требовать действий от оператора.

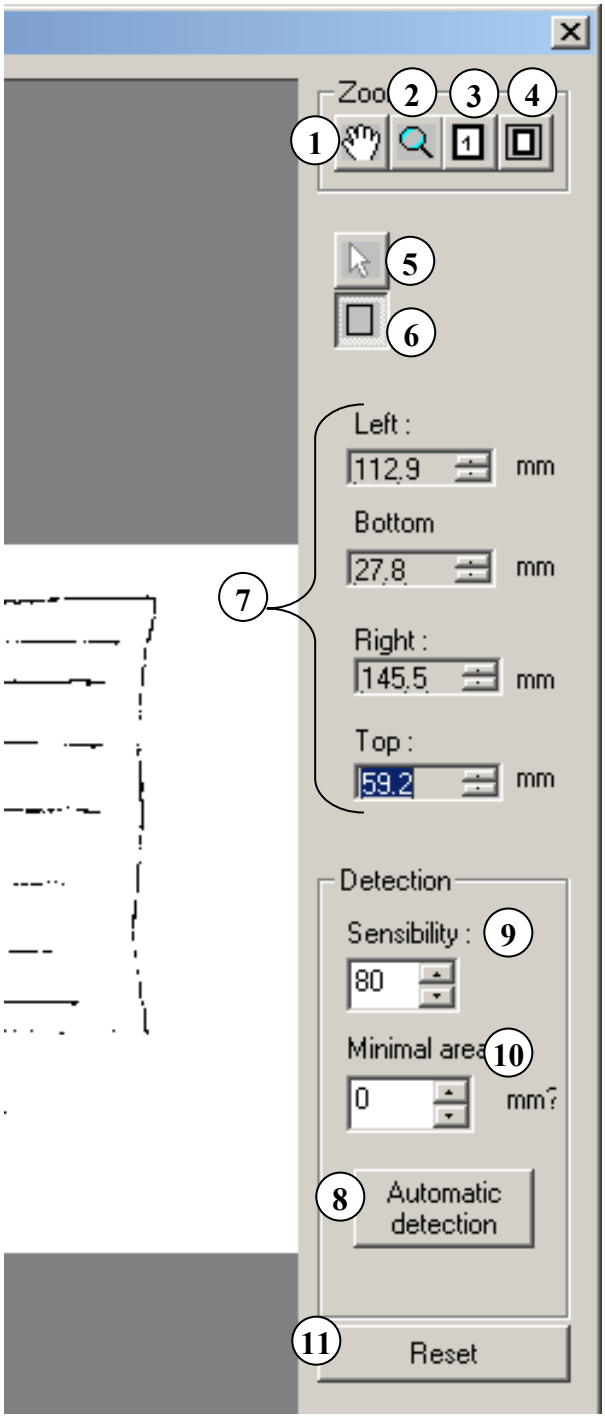
Рассмотрим какие возможности предлагает данное диалоговое окно. Итак, оно позволяет:

- создать новую зону;
- изменить положение границ зоны;
- удалить существующую зону.

Для этого в нем можно изменять размер (масштаб) изображения, сдвигать его в окне, выбирать зоны для редактирования, рисовать прямоугольники без создания зоны (проверяя границы зон) и т.д.

К сожалению, это, наверное, единственное в Book Restorer диалоговое окно, не являющееся интуитивно понятным, и я до написания этой статьи некоторых возможностей программы не видел. Дело в том, что это окно функционирует как бы в двух режимах: режиме добавления зоны и режиме редактирования зоны; и реагирует в них на действия пользователя по разному.

Диалоговое окно редактирования зон вручную содержит следующие управляющие элементы:



1) перетаскивание - позволяет двигать изображение, если оно превышает размеры окна

2) увеличение - увеличивает масштаб изображения (при зажатой клавише Ctrl - уменьшает)

3) отображение рисунка в реальном масштабе один к одному (1:1)

4) отображение рисунка на все окно

5) режим выделения существующих зон и изменения границ зон

6) режим создания зон

7) индикация расстояния от левой и правой границ зоны до левого края страницы, а также между нижней и верхней границей и нижним краем страницы (краем изображения)

8) запуск автоматического детектирования зон на изображении

9, 10) параметры автоматического детектирования: чувствительность, минимальная зона

11) удаление всех определенных зон

Как уже говорилось, если в начальном диалоговом окне были выбраны и автоматическое определение и ручное редактирование, то наше изображение откроется в данном окне с уже распознанными зонами. Если же было выбрано только ручное редактирование, то изображение откроется «как есть». Однако их данного окна можно запустить процесс детектирования (8), посмотреть сколько зон определится, и если нас это не устраивает - нажать сброс (11) и изменить параметры детектирования (9, 10). Либо вообще (при небольшом числе фактически имеющихся зон) попытаться определить все зоны вручную.

Основные процедуры ручного редактирования зон осуществляются в данном окне следующим образом:

А) проверка зоны

1) выбрать режим выделения зон (кнопка - «указатель»);
2) при зажатой левой клавише мыши на изображении будет рисоваться зеленый прямоугольник, который после отпускания клавиши будет исчезать. Позволяет проверить что именно захватывает зона.

Б) создание зоны

1) выбрать режим создания зон (кнопка - «прямоугольник»);
2) при зажатой левой клавише мыши перемещение курсора по экрану создает зону (зеленый прямоугольник с засечками);
3) клик вне зоны - прямоугольник зоны станет красным.

В) изменение размера существующей зоны

1) выбрать режим выделения зон (кнопка - «указатель»);
2) кликнуть левой клавишей мыши по интересующей нас зоне - прямоугольник из красного превратится в зеленый и на нем будут выделены засечками углы и середины сторон;
3) перемещая мышкой эти засечки, можно изменить размер зоны;
4) клик мышкой за пределами зоны - она станет красной

Г) удаление зоны

1) выбрать режим выделения зон (кнопка - «указатель»);
2) кликнуть по зоне левой клавишей мыши - (она выделиться и станет зеленой);
3) при зажатой клавише (Ctrl) можно выделить и другие зоны;
4) нажатие клавиши Delete на клавиатуре удалит все выделенные зоны

Двойной клик правой клавишей мыши на изображении позволяет переключиться между режимами просмотра (режимом выделения зон) и «редактирования» (режимом создания зон) без нажатия на соответствующие кнопки.

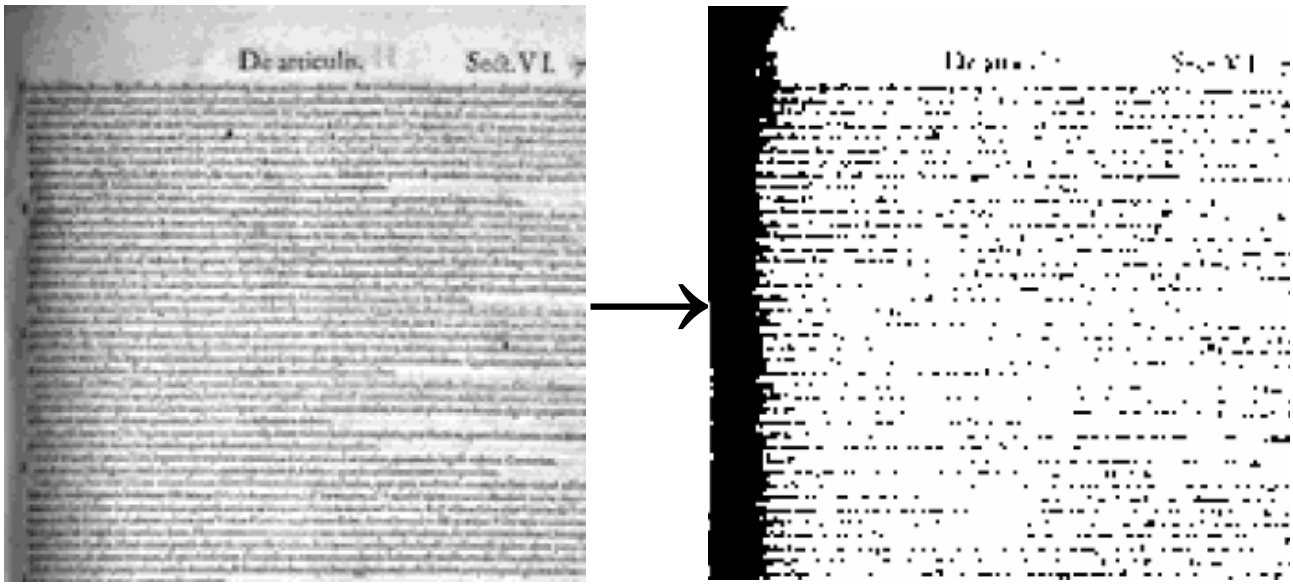
Если планируется использовать средства автоматизации Book Restorer для применения определения зон и раздельной обработки текста и изображений, то процесс необходимо осуществлять в две стадии. Первая стадия - применение определения зон ко всем страницам книги. Вторая стадия - программирование и применение к отдельным выделенным (детектированным) зонам отдельных процессов обработки.

3.3. Бинаризация

Процедура бинаризации преобразовывает серое или цветное изображение в двухцветное - черно-белое. При сканировании старых документов из-за старения бумаги, наличия теней и разнообразия текстур в результате получается изображение текста на неоднородном сером фоне. Поскольку читать удобнее черно-белые тексты, этот серый фон желательно удалить.

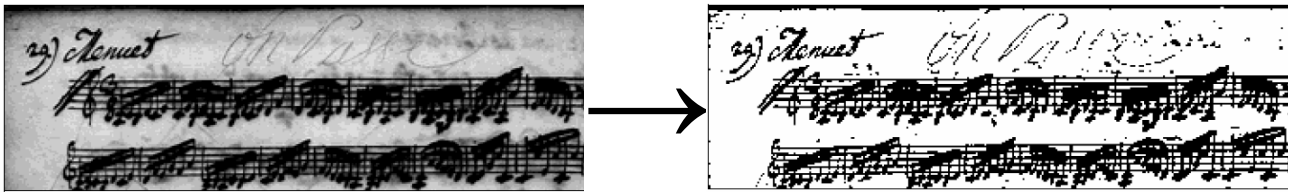
Бинаризация поэтому, наверное, самая часто используемая обработка. Помимо того, что она улучшает восприятие текста, она к тому же ведет к сокращению используемого места на жестком диске, т.к. уменьшает объем информации об изображении.

Алгоритм бинаризации в Book Restorer - локальный, то есть он не сравнивает каждый пиксель изображения с единым (глобальным) пороговым значением. Вместо этого определяются зоны с разной степенью освещенности и каждый пиксель сравнивается с индивидуальным пороговым значением. Это позволяет не допустить ошибок бинаризации, возникающих при обработке неравномерно освещенного изображения с глобальным пороговым значением:



ошибка глобальной бинаризации при неравномерной освещенности

Алгоритм также включает в себя операции фильтрации и «распространения» (propagation). После процесса бинаризации можно выполнить отдельно процесс обработки «фильтрация», чтобы убрать паразитный мусор. Такой алгоритм, как утверждает компания i2S, позволяет при необходимости сохранить детали изображения, которые в других программах исчезли бы:



сохранение карандашной пометки при бинаризации в Book Restorer

В результате пользователь программы при применении процесса бинаризации должен задать два простых и понятных параметра обработки:

1) «детали» (details) - уменьшение параметра ведет к большей фильтрации, и наоборот. Чем больше деталей, тем больше мусора, однако слишком большая фильтрация может привести к исчезновению нужных элементов изображения (слабоконтрастных букв, точек над ё);

2) «глубина» (depth) - регулирует толщину линий - слишком большое значение параметра может привести к появлению «разводов», слишком маленькое - разрывов в буквах.

разводы

разрывы

При вызове процедуры обработки откроется диалоговое окно, вид которого будет зависеть от цветности исходного изображения. Если изображение черно-белое, то окна предпросмотра не будет (и ничего процесс сделать с изображением не сможет), а если исходное изображение цветное или серое, то диалоговое окно откроется с окошком предпросмотра.

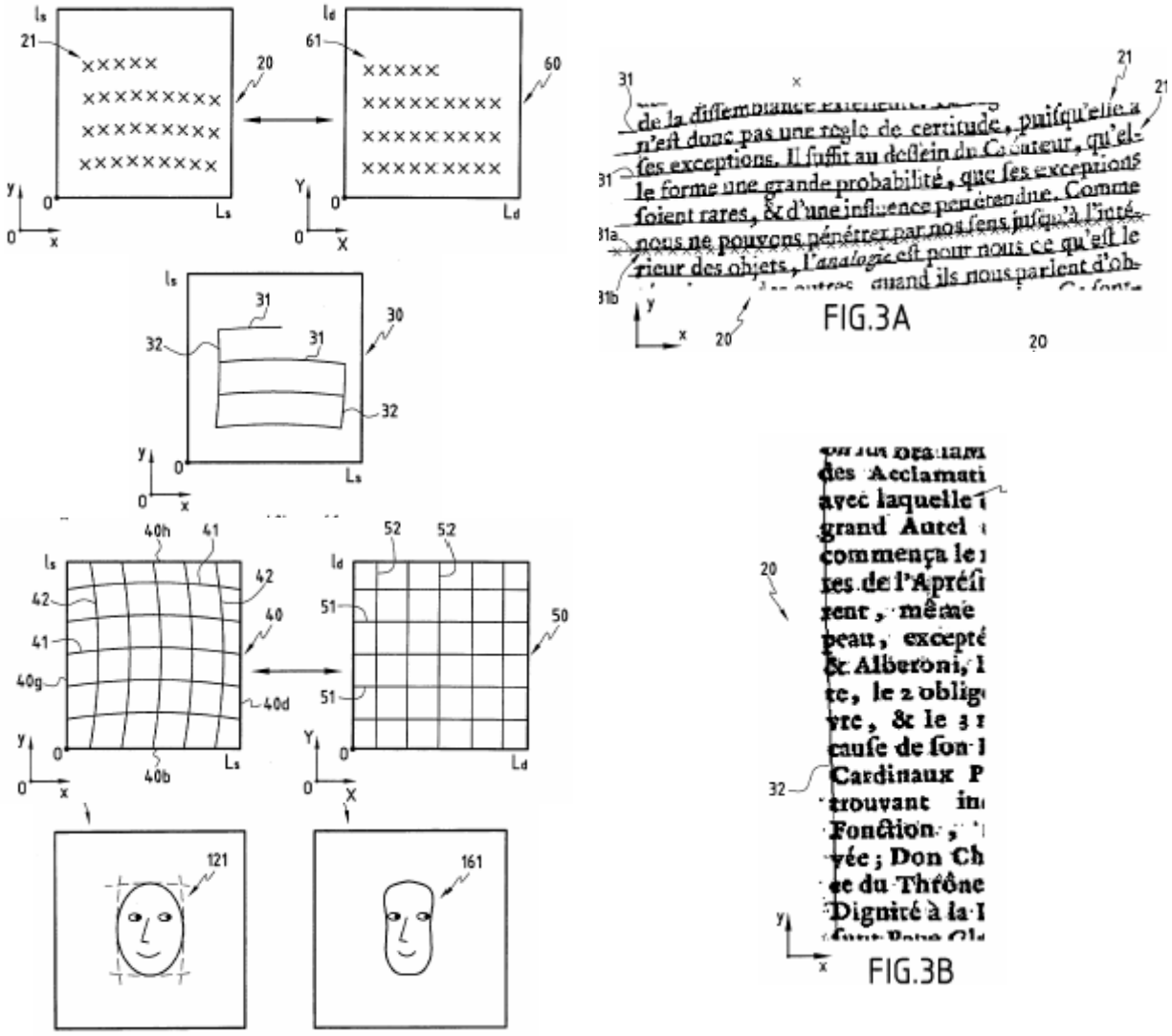
При нажатой галочке «*Quick preview*» в нем будет показываться результат применения выбранных параметров. Скроллбарами можно передвинуть предпросмотр на более важный или содержащий чувствительные элементы участок. Однако, результат который показывается в этом окошке может отличаться от конечного итога.

Нажатие на кнопку «Preview» применит выбранные параметры процесса ко всему изображению, которое видно в зоне страницы. Если не нажимать потом кнопку «ОК», то оно выполнено не будет.

Ползунки, устанавливающие параметры, двигаются как курсором так и колесиком мышки (это позволяет их точнее позиционировать). Обычно, я двигаю ползунок «детали» максимально влево. Потом двигаю ползунок «глубина» вправо, пока результат не начинает ухудшаться. Потом смещаю оба ползунка поочередно на одно деление вправо и смотрю, как меняется результат, пока не найду оптимальный вариант. По мнению разработчиков Book Restorer, бинаризацию имеет смысл делать после всех остальных процессов обработки

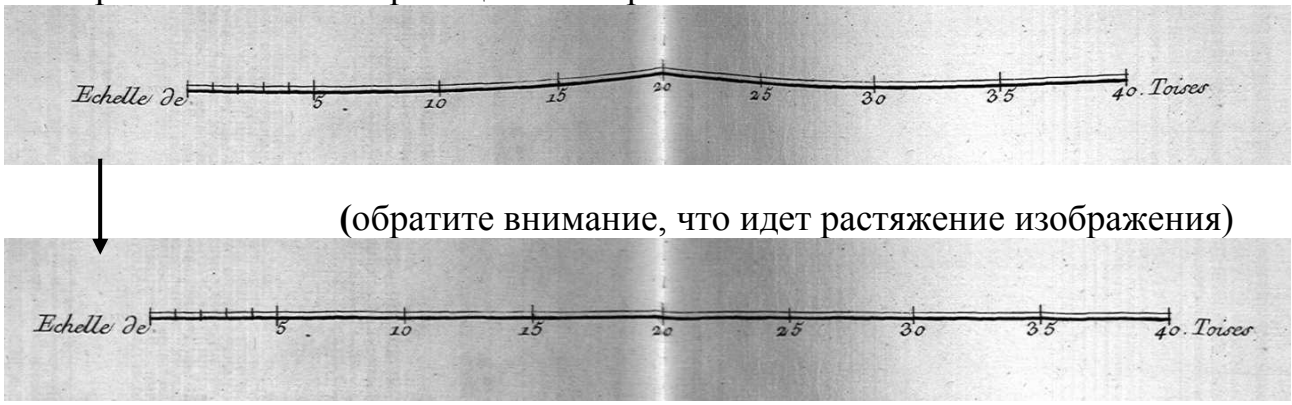
3.4. Геометрическая коррекция

Геометрическая коррекция исправляет искажения, вызванные изогнутостью страницы книги (кривизна строк), а также измятостью бумаги (пляска букв - их расположение выше и ниже линии строки):



Рисунки из патента компании i2S на алгоритм геометрической коррекции

Кроме того, данная процедура коррекции путем восстановления формы страницы в трехмерном пространстве позволяет исправить возникающее при сканировании сжатие страниц возле корешка книги:



(обратите внимание, что идет растяжение изображения)

Как указывает компания i2S в своём рекламном журнале DigiBook Mag, при сканировании возникают дефекты, связанные оригиналом (качество старой бумаги, замятия), и дефекты связанные с процессом сканирования (расположение оригинала, фотоэффекты), требующие геометрической коррекции искривлений изображения.

Использовавшиеся ранее методы коррекции заключались, как правило, в однонаправленной трансформации изображения по вертикали или по горизонтали. Также, как правило, использовался только один тип информации для оценки необходимой коррекции (либо линии текста, либо границы страницы, вертикальные и горизонтальные линии на рисунке, края абзацев и т.д.). Такой метод имеет ограничения при поврежденных краях страницы, наличии строк текста, написанного от руки, повернутого текста.

Для Book Restorer был разработан и запатентован новый алгоритм, который учитывает информацию из всех возможных источников, что приводит к потрясающим результатам в обработке отсканированных изображений.

1) исходное изображение

2) коррекция на основе краёв листа

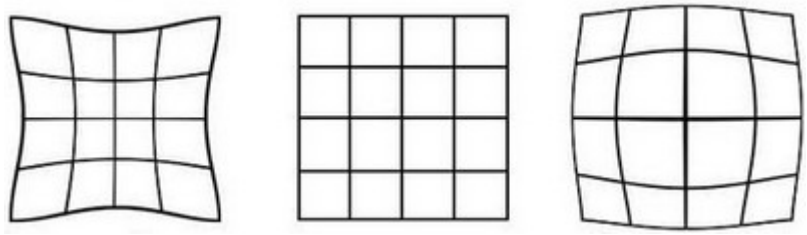
3) коррекция на основе строк текста

4) коррекция Book Restorer

Если сравнить приведенные выше изображения, то видно, что коррекция по краям листа не привела к выравниванию строк, а коррекция по строкам сделала все вертикальные лини наклонными. Коррекция же Book Restorer учла изгибы страницы, её наклон, локальные искривления и сохранила естественное восприятие изображений.

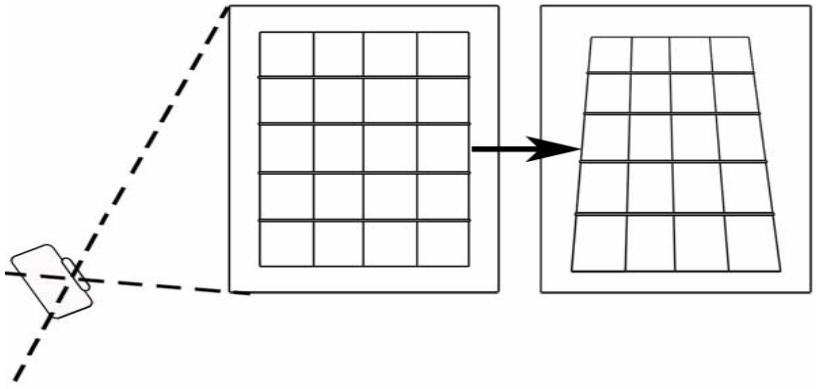
Такая универсальная коррекция (которая включает выравнивание по краям колонок текста), позволяет устранять искажения, связанные не только с недостатками документа, но и искажения, связанные с процессом фотографирования книги - «сканирования» на планетарных сканерах, поставкой которых также занимается i2S - компания-разработчик Book Restorer. С фотографированием, как известно, связаны два основных искажения изображений страниц книги - это дисторсия и искажение перспективы:

дисторсия возникает из-за различного линейного увеличения в оптической системе (объективе) разных частей изображения, в результате которого прямые линии (за исключением лежащих в одной плоскости с оптической осью) изображаются кривыми:



подушкообразная дисторсия бочкообразная дисторсия

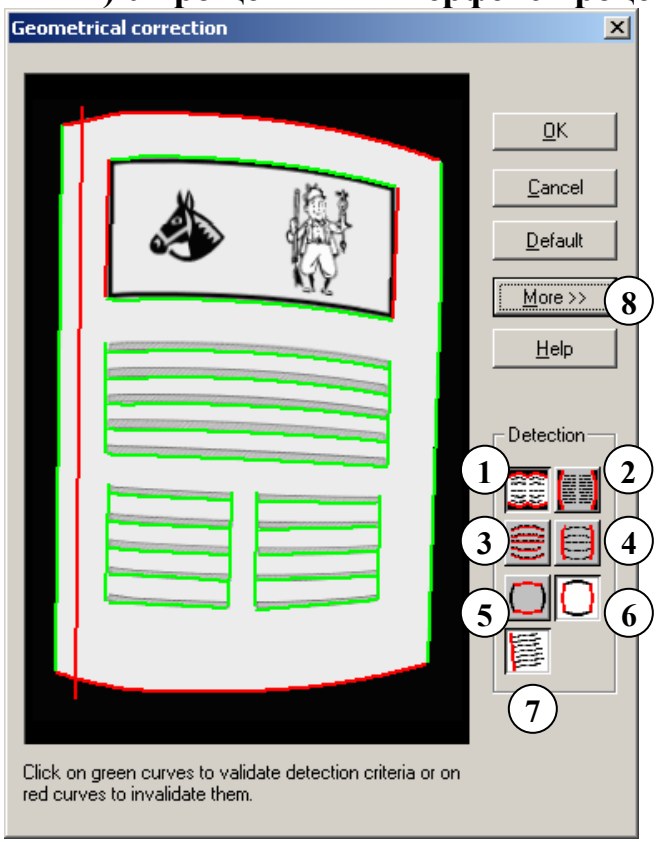
искажение перспективы возникает из-за несовпадения плоскости съемки и плоскости изображения (наклона фотоаппарата)



Интерфейс процесса геометрической коррекции в программе Book Restorer имеет два представления: сокращенное (к сожалению, это для лохов) и расширенное (которое и представляет для нас интерес). Рассмотрим сначала сокращенное представление интерфейса процесса. При вызове процесса открывается окно, которое позволяет выбрать любой набор из семи доступных детекторов, а также вызвать расширенный интерфейс процесса.

На приведённом ниже рисунке обозначены указанные детекторы. При использовании же расширенного интерфейса процесса возможным становится не только выбор используемых детекторов, но и назначение их параметров (обозначены на следующем рисунке).

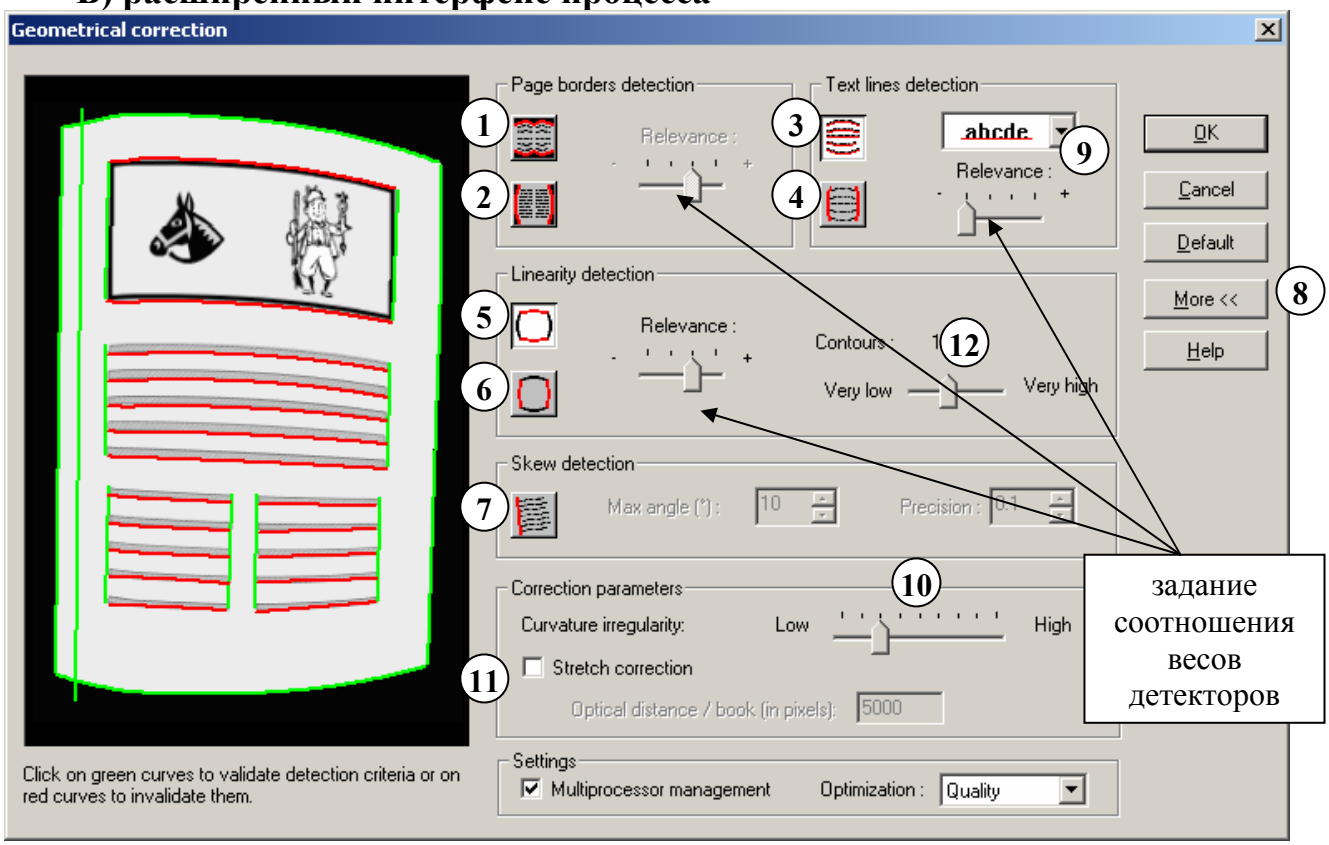
А) Упрощенный интерфейс процесса геометрической коррекции



В расширенном интерфейсе кнопки те же самые, что и в упрощенном. Помимо нажатия на кнопки включить детекторы можно и кликая по условной схеме страницы. Если детектор включен, то соответствующий элемент изображения становится красным (выключен - зеленым).

1) детектор верхней и нижней границы страницы
2) детектор левой и правой границы страницы
3) детектор линий строк
4) детектор краёв столбца текста
5) детектор горизонтальных линий
6) детектор вертикальных линий
7) детектор наклона
8) переключение интерфейса
9) выбор способа определения линии строки
10) выбор интенсивности локальных нерегулярностей кривизны
11) включение исправления сжатия (у корешка книги) - «растягивание»
12) выбор порогового значения для разделения текста и фона

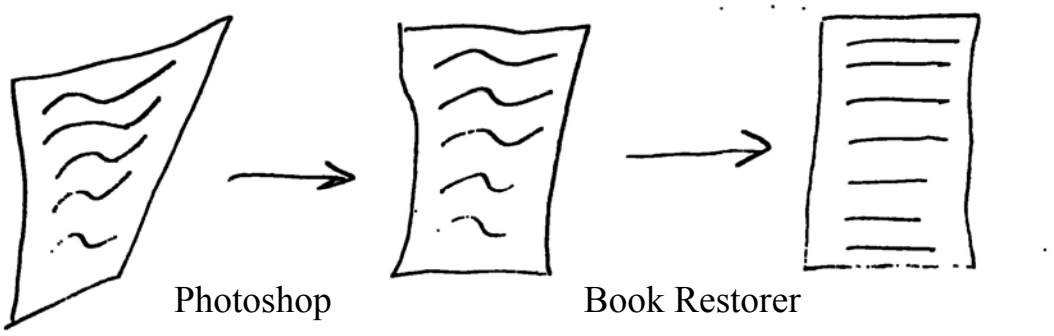
Б) расширенный интерфейс процесса



Итак, как же следует пользоваться данным модулем. Прежде всего, следует учесть общий принцип работы с Book Restorer - выбрать наиболее типичную страницу, подобрать для нее обработку. Потом применить такую обработку сразу ко всем страницам книги. А вот потом уже отслеживать какие страницы не получились, и обрабатывать их отдельно вручную, подбирая для каждой свои параметры.

Процесс геометрической коррекции иногда дает неудовлетворительные результаты - сильно искажает страницу. Если применять стандартные параметры для всех страниц, то процентов 10 окажется побитыми. Это нестрашно. Практика показывает, что даже малейшее изменение параметров обработки может избавить от таких «странных» эффектов как изломы строк, растяжения по вертикали, неправильное детектирование края текста.

Кроме того, хочется отметить два важных момента. Во-первых, у процедуры существует ограничение на степень искажений. То есть, если на странице есть сильно искривленный блок текста (иногда такие страницы выявляются только после примененной обработки - это те, что не получились), то лучше подровнять его в Photoshop - сделать страницу более прямоугольнообразной и только тогда применять к ней процедуру коррекции.



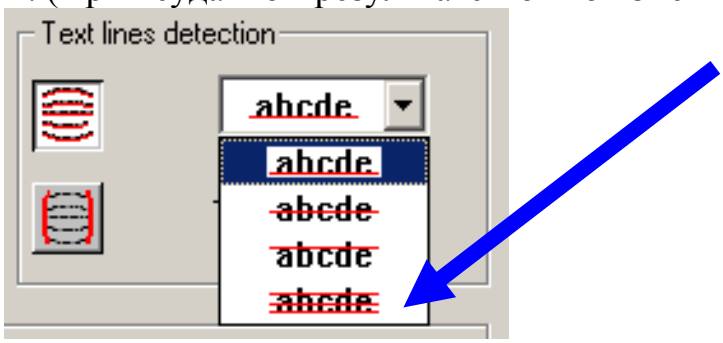
Во-вторых, если на странице есть блоки текста с разным характером искажений или разной природы (у одного есть край столбца текста - у другого нет; один наклонен в одну сторону - другой в другую), то такие блоки следует распознать как разные зоны (процесс Areas detection в списке инструментов реставрации Book Restorer) и применять процедуру геометрической коррекции к таким зонам по отдельности. Точно также, если текстовый блок занимает лишь небольшую часть страницы (например, последняя страница какой-либо главы), то результат обработки будет сильно лучше, если геометрическая коррекция будет применена не ко всей странице целиком, а именно к зоне с этим текстом.

До того как начать выбирать параметры процесса, следует определить имеющиеся на странице элементы, чтобы потом правильно включить детекторы. Потому что если соответствующий элемент отсутствует, то процедура скорее внесет искажение, чем что-либо исправит. Здесь, нас конечно прежде всего интересует наличие вертикальных/горизонтальных линий и наличие по бокам или сверху и снизу страницы черных краев (как правило от крышки сканера там, где книги на стекле уже не было)[2].

[2] Это, кстати, означает, что процесс обрезания лишнего следует делать ПОСЛЕ геометрической коррекции.

1) Включим детектор строк текста (3), который опирается на имеющие достаточную длину строки. Несмотря на то, что в справке указано, что поддерживаются только латинские буквы, кириллица тоже неплохо выравнивается. (Можно, впрочем, представить пример - какие-нибудь иллюстрации в рамках на всю страницу - в котором строк текста нет, а прямые линии есть. Тогда, конечно, включать этот детектор не следует.)

2) Выберем способ определения линии строки (9) - на какую ее часть (верхнюю, среднюю или нижнюю) будет ориентироваться детектор. Я обычно выбираю самую последнюю опцию, в которой выравнивание идет и по верхней и по нижней кромке строки. (При неудачном результате можно изменить.)

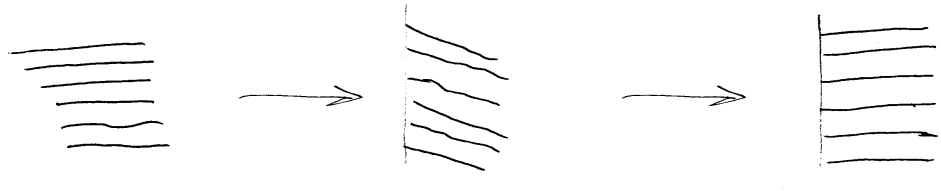


3) Включим, если это необходимо, кнопку исправления краев текста (4). Скажем так, это - не детектор. Я бы сформулировал это таким образом - если надо поправить края текста, то включим. Если не надо, то лучше не включать, вне зависимости от того, есть ли у столбца текста четкий край. Искажений от него больше. Не случайно эта кнопка расположена внутри поля «строки текста».

4) Включим нужные детекторы (1), (2), (5), (6) в зависимости от того, какие опорные элементы присутствуют на изображении, и какие мы хотим поправить, есть ли чёрные поля (рамка), возникающие вокруг страницы при сканировании, рамки вокруг рисунков, или проведенные между колонками или параграфами черты. Если этих элементов нет, то включать детекторы не следует.

5) Выберем веса детекторов (их релевантность). Релевантность отражает влияние на коррекцию выбранного критерия. Детектор с низкой релевантностью дает меньшее влияние, чем детектор с большей. Два детектора с низкой релевантностью дадут тот же результат, что и те же два детектора с высокой релевантностью. Я обычно ставлю строки на максимум, остальные (если они включены) - на минимум. При неудачном результате изменяю.

6) Включим исправление наклона (7) Практика показала следующее: исправление наклона в принципе (в частности, в других программах, из которых я и перекидываю промежуточные результаты обработки в Book restorer) можно делать по правому/левому краю текста, по центру, по достижению горизонтальности строк и т д. Процедура геометрической коррекции Book Restorer, по-видимому, оптимизирована под текст, выровненный по левой стороне, в то время как другие программы (например, Scan Tailor) равняют по-другому. Если её не включить, то поедут края строк, а если включить, то процедура сначала скосит строки, а потом их выправит:



При неудачном результате (изломы, растяжения) опцию можно отключить. При определении наклона возможно задание тех же параметров, что и в процессе обработки «Исправление наклона» (Deskew) максимальный угол, точность.

7) Выключим курватуру (curvature irregularity). Соответствующий ползунок (10) по умолчанию стоит на третьей позиции из девяти возможных. Этот параметр определяет регулярность искажений. Как указано в справке: «*Процесс геометрической коррекции использует глобальную информацию о кривизне. Для мятой или поврежденной бумаги выберите низкую нерегулярность, чтобы ограничить возмущения детектора»*. Мой опыт показывает следующее: данный параметр может помочь подровнять «пляшущие буквы». В тоже время если буквы не пляшут, то надо вообще выключать - иначе они как раз в разнос и пойдут. То есть если бумага не помята, то в ноль. А если помята, то включить. Но помня, что написано в файле справки, на максимум я параметр не ставлю.

8) Включим в случае необходимости растяжение изображения (11) у корешка книги. Это растяжение (и именно у корешка) работает даже для изображений одинарных страниц. Если его включить, то становится также доступным параметр Optical distance (оптическое расстояние) - расстояние между фотоаппаратом и книгой, необходимый для оценки перспективы. Я обычно эту коррекцию не включаю, а если и включаю, то оставляю для сканированных книг значение параметра по умолчанию (5000). Оно задается в пикселях так: $расстояние\ (pix) = расстояние\ (cm) \cdot \frac{разрешение\ (dpi)}{2{,}64}$.

Наверное, в случае фотографирования книг дома фотоаппаратом со штатива (или сканирования реально толстых книг на сканере) можно откалибровать этот параметр - заложить в книгу распечатанное изображение линейки (типа как на рисунке, приведенном в начале параграфа) и подобрать оптимальное значение.

У нас остался один параметр, который я лично никогда не изменял и практической информацией о его работе не владею. Это «контур» (12). Он влияет на детектирование прямых линий, определяет пороговое значение для разделения текста и фона. Как написано в справке Book Restorer, «*максимальное значение в 255 соответствует вариации между пиком белого и пиком черного цвета. Значение 0 соответствует отсутствию изменений яркости. Следует выбирать значение в диапазоне 64-192, чтобы только толстые элементы изображения учитывались при обработке*».

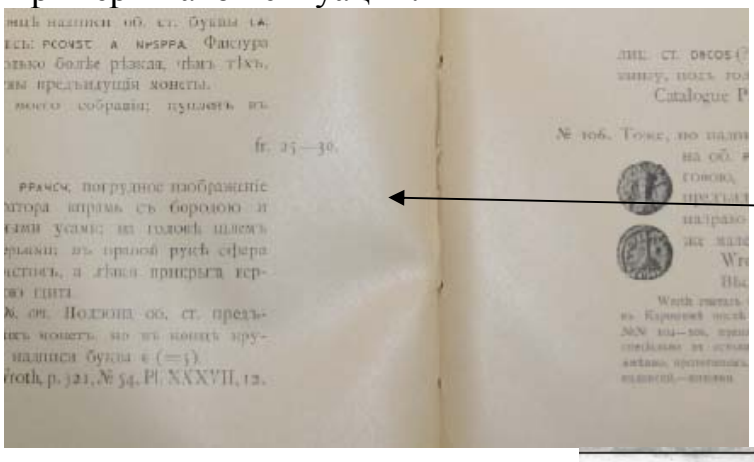
Важный замеченный нюанс: **очередность процедур коррекции влияет на результат**. То есть если коррекция строк и линий одновременно дала плохой результат, то следует попробовать исправить сначала строки потом линии. Затем сначала линии, потом строки. Также влияет изменение их релевантности.

3.5. Коррекция освещенности

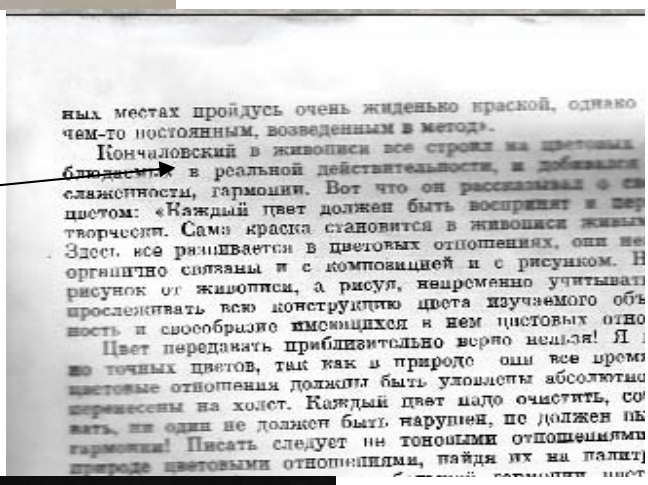
Коррекция освещенности исправляет неравномерность освещения страниц, которая возникает, как правило, в следующих случаях:

- смещение вспышки от фотографирования;
- блики от глянцевой страницы при фотовспышке;
- затемнение в зоне корешка книги;
- измятие бумаги, приводящее к варьированию расстояния между страницей и лампой сканера;
- нестабильность работы лампы сканера.

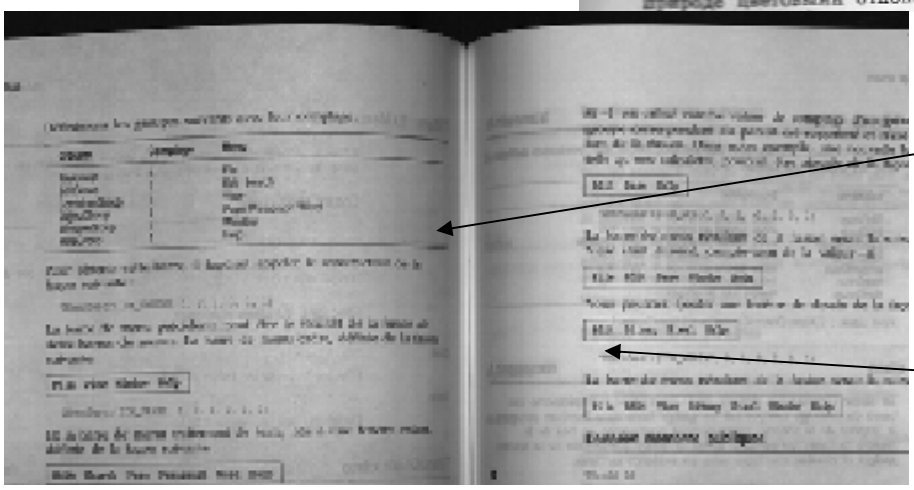
Примеры такой ситуации:



блики от вспышки
на изогнутой бумаге



бумага отходила
от CIS сканера
(размывы, пятна в освещённости)



затемнение
у корешка

нестабильность
лампы сканера
(светлые полосы)

В принципе, поскольку обычно мы переводим книги в черно-белый формат и делаем это с помощью локальной бинаризации (которая подстраивается под участки с разной освещенностью), коррекция освещенности пригождается не слишком часто. Я бы выделил три принципиально разных ситуации, когда эта процедура обработки действительно оказывается полезной:

а) такие перепады освещенности у корешка, что процесс бинаразиции дает ошибки, а бинаризация, выполненная после коррекции освещенности, выдает приемлемые результаты;

б) блики от вспышки (или другие факторы) привели к пятнам освещенности на картинках, которые в окончательной книге останутся цветными (не будут бинаризоваться),

в) неравномерная освещенность и размывы (как правило, на изображениях с использованием CIS-сканера) ставят под угрозу перемычки в буквах, прежде всего в «и» и «н».

На мой взгляд, три доступных программы для обработки сканов книг: freeware-программа Scan Kromsator; свободное программное обеспечение Scan Tailor и демонстрационная версия (TryOut) Book Restorer принципиально по разному подходят к процедуре коррекции освещенности. По-видимому, их разработчики за приоритет брали именно различные эти ситуации. По моим ощущениям, работают они следующим образом:

ST - вычитает некую нелинейную функцию интенсивности освещенности из изображения, в результате чего получаем равномерно освещенный но слишком светлый скан. Я им пользуюсь для обработки сканов страниц с картинками, а потом делаю автоопределение уровней в Фотошопе, чтобы получить более темный и с правильными цветами итог;

SK - обладает большим количеством настроек коррекции освещенности, и в принципе позволяет получить довольно разнообразные по сравнению с параметрами по умолчанию результаты, но во-первых перебрать все параметры невозможно; во-вторых, он заточен на последующую бинаризацию. Хорошо работает с корешками. В крайнем случае, зону у корешка а в нем можно выделить отдельно и бинаризовать с отдельными настройками, не как всю остальную страницу;

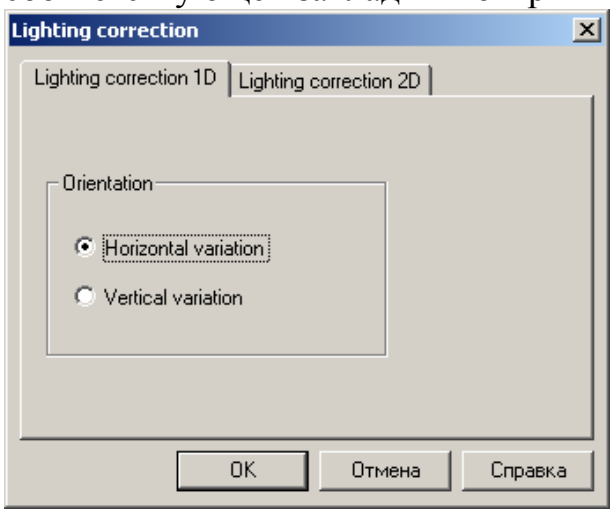
BR - как бы (по моим визуальным ощущениям и в сравнении с другими программами обработки) усредняет фон вокруг букв; то есть как бы делает сегментацию текст-фон, а потом высчитывает среднее значение для фона и делает равномерную заливку фона. Я его применяю в тех редких случаях, когда перебираю разные варианты (и в разных программах) лишь бы спасти попорченные кривым сканированием буквы. Иногда коррекция освещенности Book Restorer оказывается полезной.

Если же книга обрабатывается без привлечения третьих программ, безусловно коррекция освещенности Book Restorer окажется полезной и при других случаях.

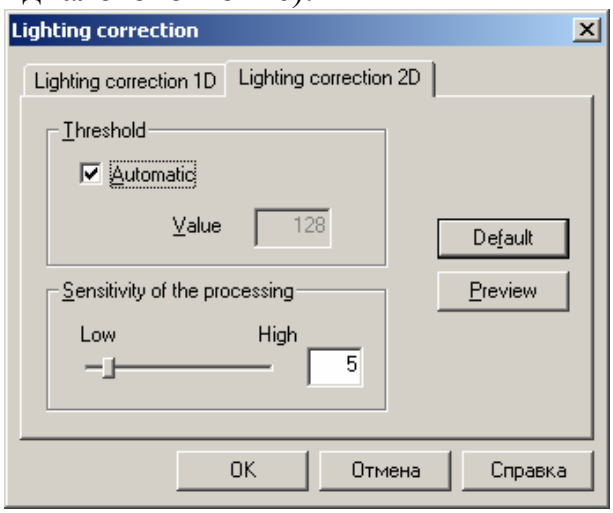
Разработчики Book Restorer выделили две принципиально различных ситуации неправильной освещенности: а) варьирование освещенности в одном измерении, например, по горизонтали (корешок книги); б) варьирование

освещенности в двух измерениях: и по горизонтали, и по вертикали из-за особенностей сканера (см. последний пример).

Соответственно при вызове соответствующего процесса обработки необходимо сделать выбор между этими вариантами (путем выбора соответствующей закладки в открывшемся диалоговом окне):



а) одномерная неравномерность



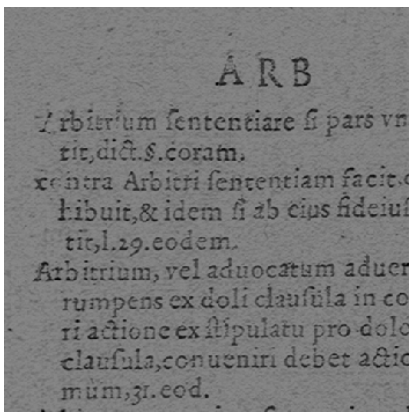
б) двумерная неравномерность

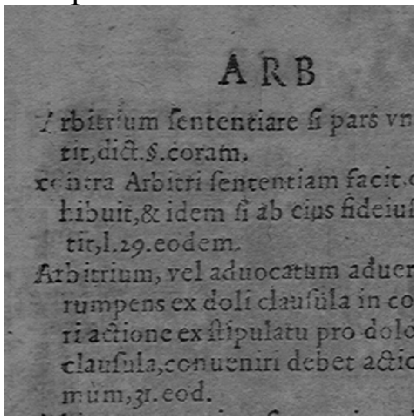
При одномерной коррекции надо выбрать по горизонтали или по вертикали варьируются светлые и темные участки изображения. Разработчики Book Restorer рекомендуют применять именно эту обработку, если на странице есть графики и фотографии. Применять ее можно и неоднократно.

При двумерной коррекции уже есть два параметра, значения которых необходимо выбрать: пороговое значение и чувствительность.

Параметр «пороговое значение» определяет среднюю яркость изображения. По умолчанию стоит галочка «автоматическое», и она будет рассчитано без участия пользователя. Задать значение вручную следует, если вокруг изображения есть черные полосы, которые впоследствии будут отрезаться. Значение варьируется между 0 и 255. Значение по умолчанию - 128.

Параметр «чувствительность» определяет игнорирование текста при коррекции освещенности. Чем больше значение параметра, тем больше элементов попадает в текст и не затрагивается обработкой:

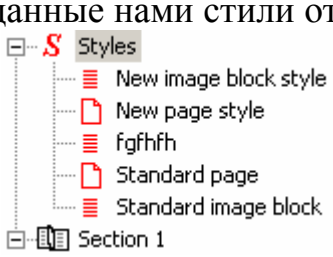




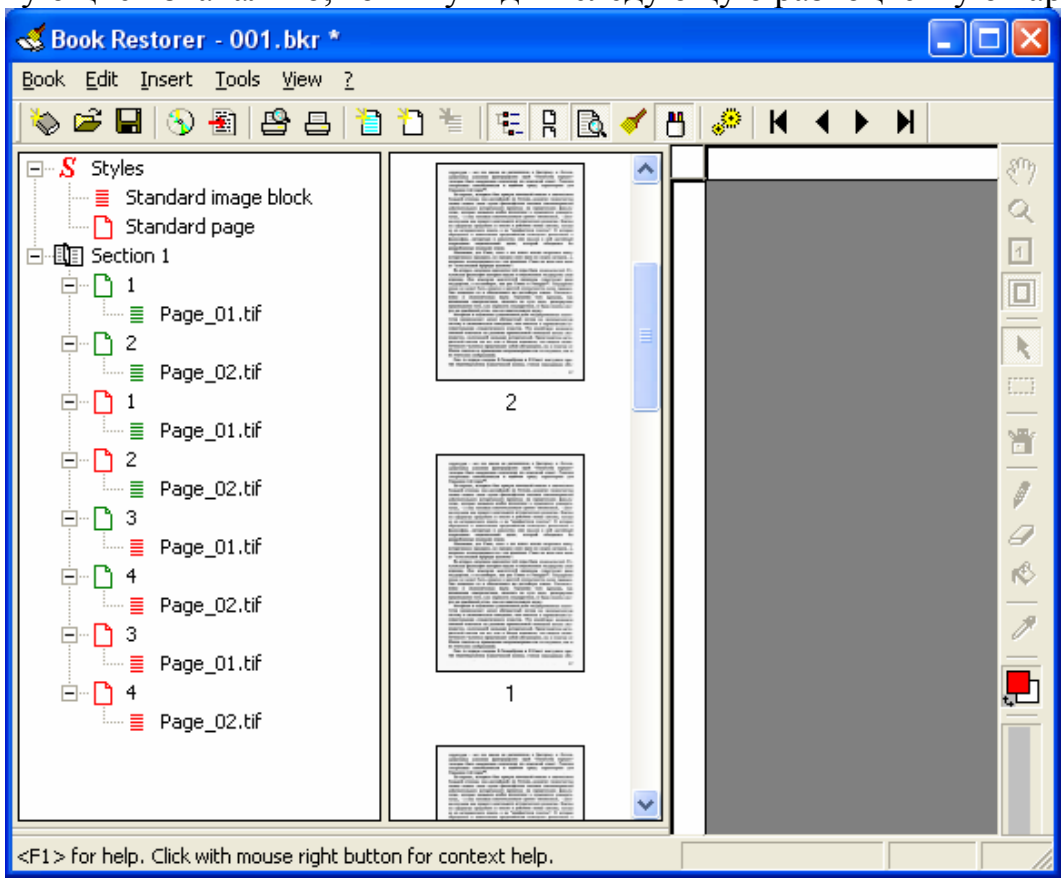
На мой взгляд, параметр чувствительность желательно ставить выше дефолтного значения (=5), но не слишком, иначе останутся грязь и пятна, как в приведенном выше примере. Высокие значения нужны и при крупном тексте.

3.6. К вопросу об использовании стилей

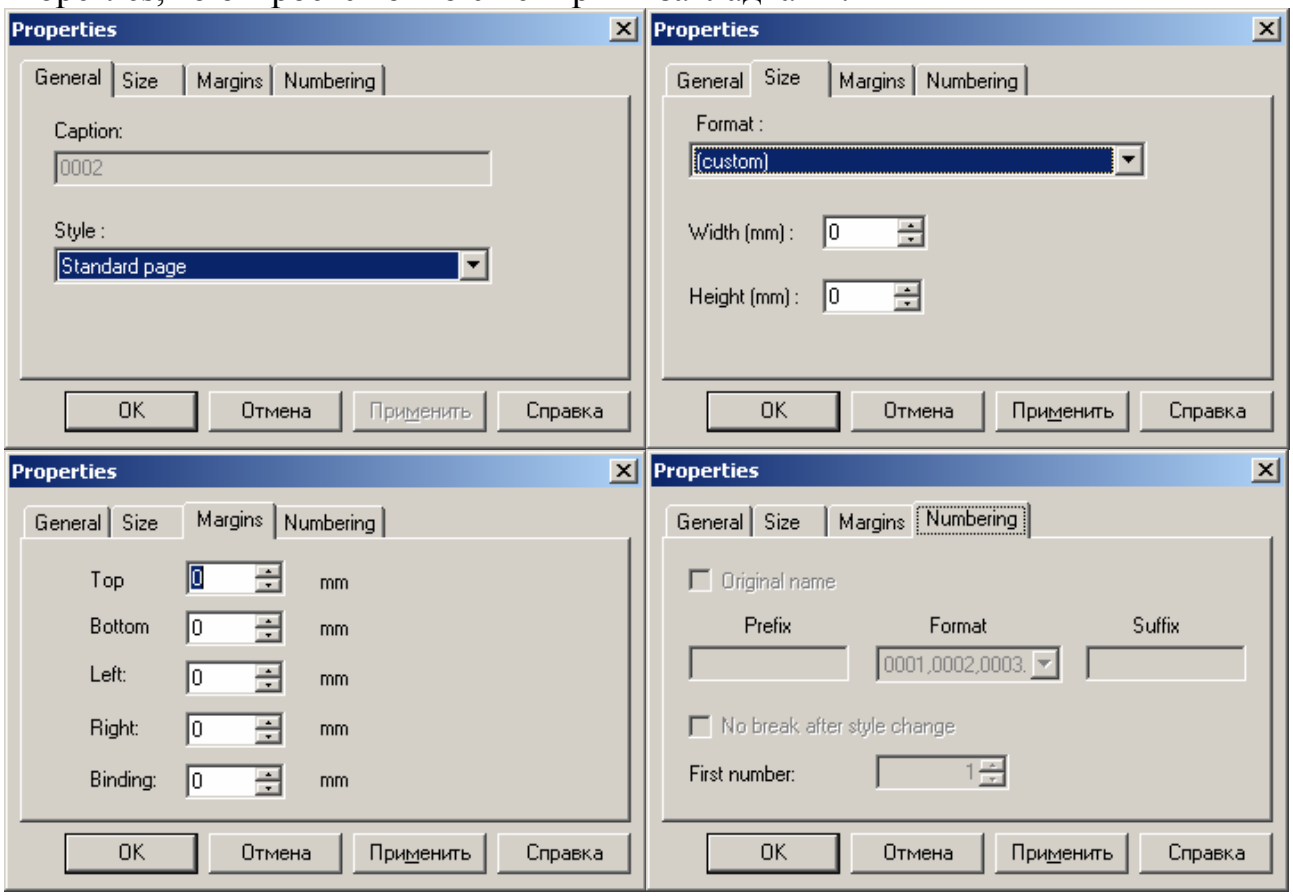
При создании книги нас спрашивают, какой шаблон (template) для книги мы выбираем. Пока нет своих шаблонов, есть два выбора «нет» и «стандартный». Шаблон - это такой файл (хранится в папке Templates в директории, в которой установлен Book Restorer), в котором хранятся скрипты (сгруппированные в наборы процессы обработки с сохраненными параметрами) и стили. Любую нашу книгу, в которой мы в процессе обработки насоздавали стилей и скриптов, можно сохранить как шаблон (Book - Save as template). Созданные нами стили отражаются в начале дерева книги в разделе Styles



Выбранный стиль отмечается в дереве книги цветовым кодированием - те стили, которые есть в дереве книги («стандартный» и наши собственные) - обозначаются красным. Если выбран вариант «без стиля» или стиль в процессе был отредактирован, то соответствующая страница или блок становятся зелеными. Так, например, если при импорте каждых двух страниц мы переберем все возможные варианты для стиля страницы и стиля блока (нет, нет; стандартный, нет; нет, стандартный; стандартный, стандартный), существующие изначально, то мы увидим следующую разноцветную картину:



Чем же различаются эти ситуации ? Стиль - это свойства страницы или блока. Если щелкнуть правой кнопкой мыши на любой странице и выбрать Properties, то откроется окно с четырьмя закладками:



На первой закладке мы можем выбрать стиль из имеющихся.

При выбранном стиле «стандарт» все остальные три окна будут неактивными. Если мы сменим его на «none» (без стиля), то включатся второе и третье окно - мы можем менять размер страницы и поля. Поля страницы на вид нашей страницы не влияют (они учитываются при определенных параметрах «опубликования). Смысл устанавливать для каждой страницы свой собственный, а не единый глобальный размер в миллиметрах я тоже не вижу. Единственная важная для нас вкладка «Нумерация страниц» остается при таком редактировании свойств недоступной. Поэтому редактирование свойств страниц (выбор «без стиля») мне кажется бесполезным.

По умолчанию, после установки Book Restorer, если ничего не редактировалось, для страниц стиль «стандарт» предполагает:

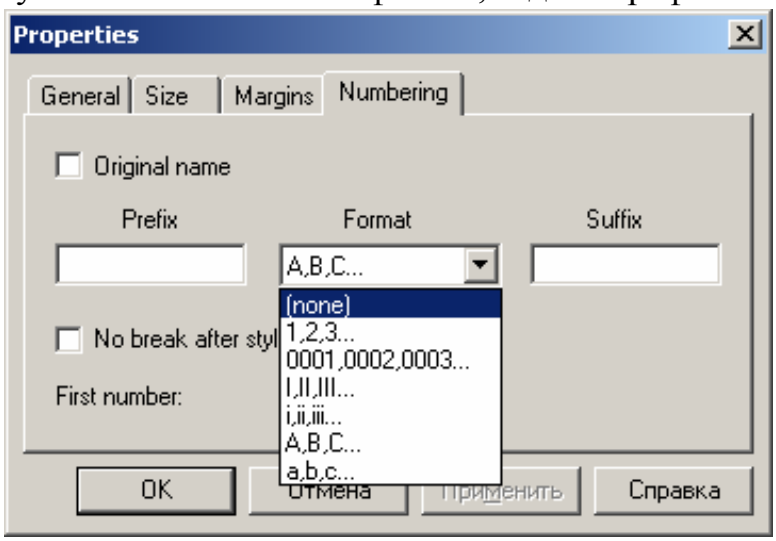
- определение размера - автоматическое (т.е. по размеру изображения);
- поля - нулевые (т.е. отсутствие полей);
- нумерацию: по типу 1,2,3, ..., 10, 11 ..., 100, 101 ...

Для того, чтобы изменить эти свойства, необходимо правой кнопки мыши кликнуть на стиль «стандарт» в разделе «стили» дерева книги, выбрать Properties, и сделать необходимые изменения. При этом откроется окно с теми же самыми четырьмя закладками. Только на первой закладке не будет возможности выбрать стиль, зато будет можно его переименовать (станет активной строка Caption); также станет активной четвертая закладка Numbering.

Прежде всего, конечно, следует изменить нумерацию, потому что она влияет в том числе на то, как будут называться итоговые файлы после их «опубликования». Для того, чтобы в других программах не варьировалась очередность файлов, правильнее нумеровать их 0001, 0002, и т.д. Если мы теперь нажнем «ОК», то при любом последующем запуске программы Book Restorer, это изменение стиля «стандарт» будет применяться.

Кстати, если импортировать файлы «без стиля», то нумерация будет идти так, как она была изначально по умолчанию: 1, 2, ... 10, ... 100.

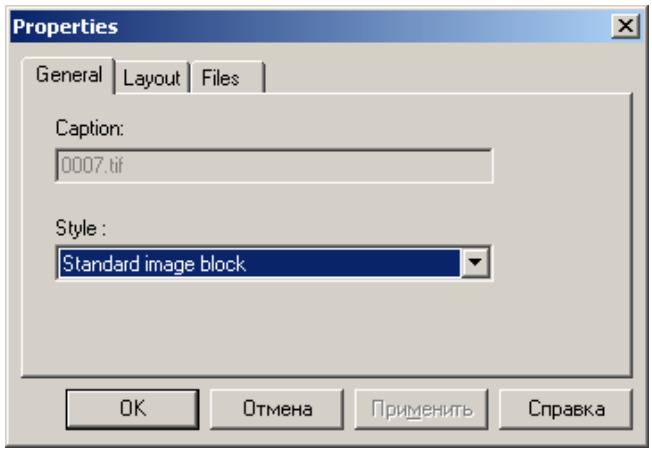
На закладке «нумерация» также можно выбрать нумерацию малыми и большими римскими цифрами, буквами латинского алфавита, задать префикс и суффикс, которые будут добавляться в начале и в конце к номеру страницы соответственно. Кроме того существует опция и «без номера»:



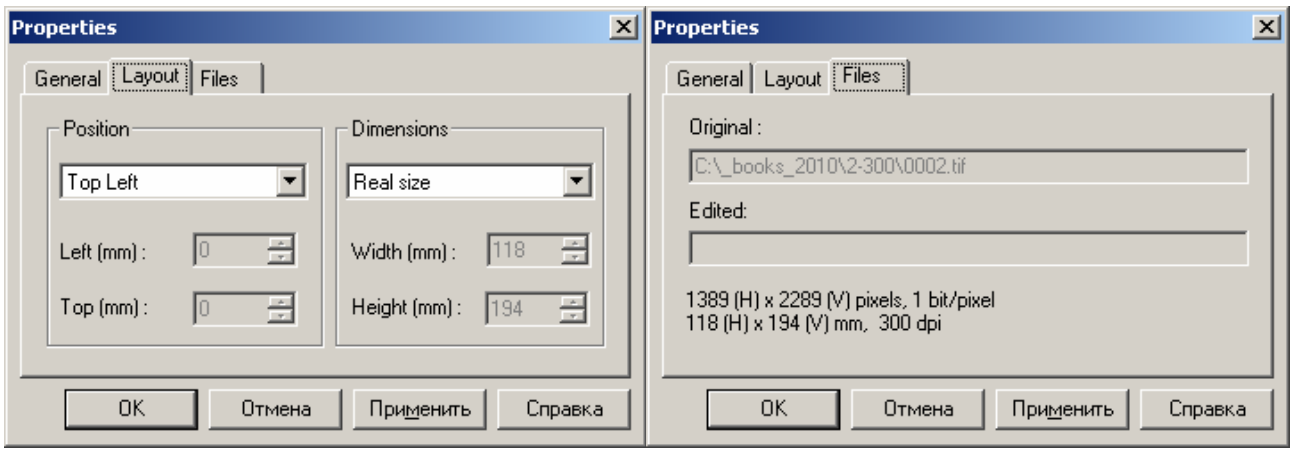
Галочка «Original name» позволяет в качестве номера (имени) страницы использовать имя файла основного изображения. Галочка в поле «*No break after style change*» позволяет сохранить прежний стиль названий и нумерации страниц при изменении стиля. В поле «*First number*» можно задать с какой цифры (буквы) начнется нумерация.

Некоторый смысл имеет также задание фиксированного размера страницы. Так, для сканированных журналов из-за рисунков часто невозможно точно определить поля. Поэтому можно задать единый размер страниц и двигать отсканированное изображение внутри этого шаблона страницы. Такое изменение, впрочем, лучше применять не в стандартном а в своем собственном стиле, создать который можно из меню *Insert - New - Page style*.

Если щелкнуть правой кнопкой мышки на любом блоке (изображении) и выбрать свойства, то мы увидим окно с тремя закладками. На первой (general) точно также как для свойств страницы можно увидеть поля «Caption» и выбор стиля.

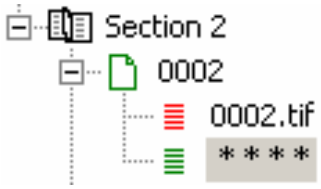


Две других закладки - это макет (layout) и файлы (files):



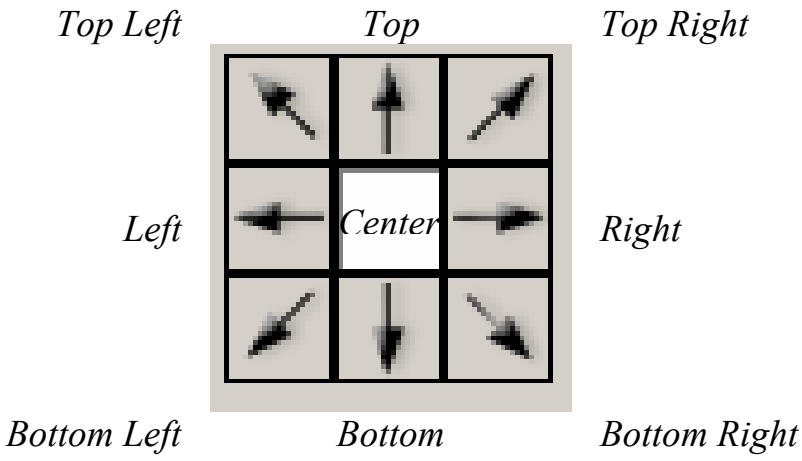
На закладке «файлы» можно увидеть, где находится промежуточное изображение, если к нему уже применялся процесс обработки, а также свойства обработанного изображения. (Если изображение не обрабатывалось, то на закладке указываются свойства исходного.)

Если на первой закладке General выбрать стиль «none», то включится вторая закладка, на которой можно выбрать положение блока изображения на странице и его физические размеры. Изменение физических размеров растягивает или ужимает нашу картинку (блок) с сохранением пропорций. При каждом изменении размера детектированной зоны к ней в дереве книги дорисовывается звездочка:

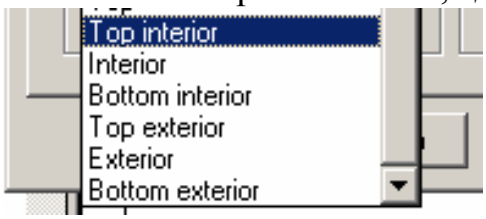


При изменении положения блока в качестве первого варианта существует возможность выбрать опцию «custom» - в этом случае включится расстояние в миллиметрах, задаваемое от левой и от верхней границы страницы. Оно может быть и отрицательным, тогда часть блока как бы уйдет за край страницы и показываться не будет. Это задание положения вручную удобно для обработки сканов журналов.

Второй вариант, это выбрать одно из девяти стандартных положений блока по отношению к краям страницы:



Третий вариант - учитывает положение блока по отношению к корешку страницы. Опция «Interior» означает «у корешка» и на нечётных страницах блок располагается у левой границы страницы, на чётных - у правой. Опция «Exterior» означает «у обреза книги», и соответственно наоборот (на нечётных - справа, на чётных - слева). Кроме смещения «к» или «от» корешка книги, здесь также задается смещение вверх, вниз или по центру, поэтому всего существует шесть вариантов выбора:



По умолчанию, после установки Book Restorer, если ничего не редактировалось, для блоков стиль «стандарт» предполагает:

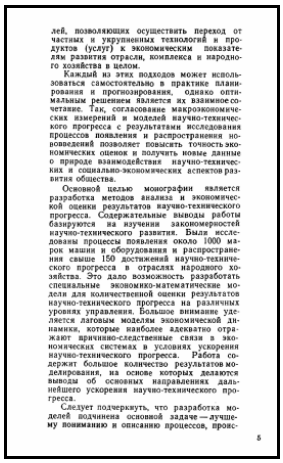
- положение - «по центру»
- размеры - «реальный размер»

Эти же установки будут выбраны и в том случае, если мы импортируем изображения с выбором «без стиля», только в этом случае эта вторая закладка сразу включена и эти свойства сразу доступны для изменения.

Хотелось бы обратить внимание, что положение блока относительно страницы определяется по краям страницы и указанные в свойствах страницы поля на это положение не влияют. Возникает вопрос, а на что же они влияют ?

Единственная ситуация, когда мне удалось обнаружить влияние полей страниц на итоговый результат - это когда при «опубликовании» выбрана опция «Activate page setup». При этом конечный результат в том числе зависит и от свойств блока, и от того, была ли выбрана опция «custom» или «automatic» в свойствах страницы (хотя эта опция не выделяется как выбранная, но почему-то это оказывает влияние на конечный размер страницы). По большому счету, результат является непредсказуемым, и лучше это не использовать[3]:

размер страницы - custom
Activate page setup - выключена
свойства блока - Top Left
поля - заданы


лей, позволяющих осуществить переход от частных и укрупненных технологий и продуктов (услуг) к экономическим показателям развития отрасли, комплекса и народного хозяйства в целом.

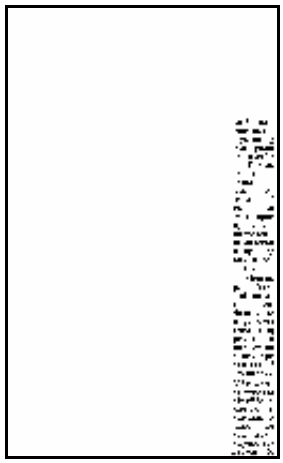
Каждый из этих подходов может использоваться самостоятельно в практике планирования и прогнозирования, однако оптимальным решением является их взаимное сочетание. Так, согласование макроэкономических измерений и моделей научно-технического прогресса с результатами исследования процессов появления и распространения новшеств позволяет повысить точность экономических оценок и получить новые данные о природе взаимодействия научно-технических и социально-экономических аспектов развития общества.

Основной целью монографии является разработка методов анализа и экономической оценки результатов научно-технического прогресса. Содержательные выводы работы базируются на изучении закономерностей научно-технического развития. Были исследованы процессы появления около 1000 марок машин и оборудования и распространения свыше 150 достижений научно-технического прогресса в отраслях народного хозяйства. Это дало возможность разработать специальные экономико-математические модели для количественной оценки результатов научно-технического прогресса на различных уровнях управления. Большое внимание уделяется макровым моделям экономической динамики, которые наиболее адекватно отражают причинно-следственные связи в экономических системах в условиях ускорения научно-технического прогресса. Работа содержит большое количество результатов моделирования, на основе которых делаются выводы об основных направлениях дальнейшего ускорения научно-технического прогресса.

Следует подчеркнуть, что разработка моделей подчинена основной задаче — лучшему пониманию и описанию процессов, проис-

5

размер страницы - custom
Activate page setup - включена
свойства блока - Top Left
поля - заданы



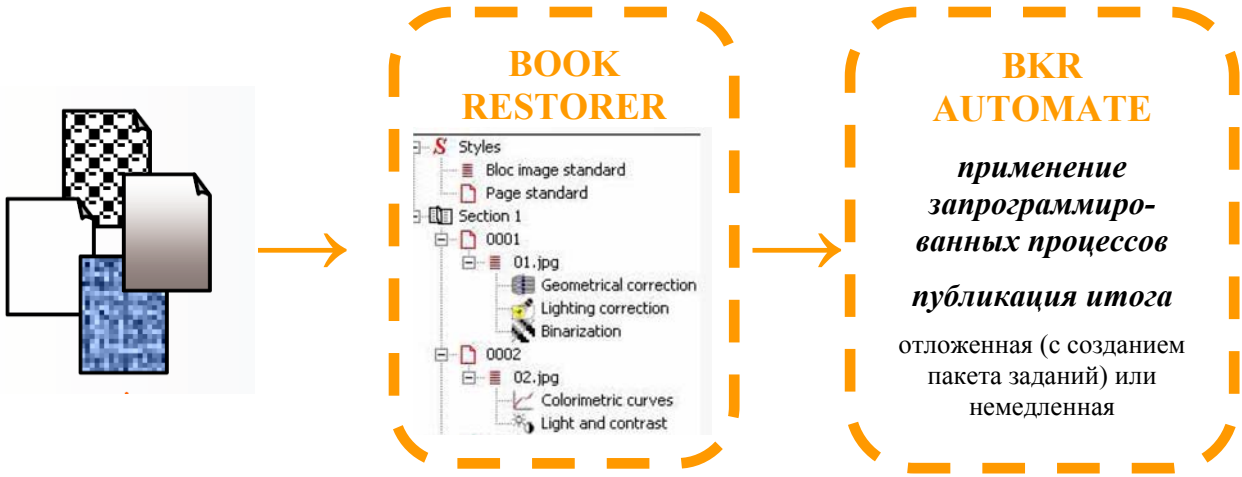
[3] В руководстве пользователя говорится, что этой опцией следует пользоваться при подготовке файлов к печати на принтере (и соответственно задании размеров печатных страниц A3, A4 и т.д., превышающих размер нашей страницы), а не при обработке изображений.

5. Работа дополнительных модулей

5.1 BKR Automation - автоматизация обработки

Модуль автоматизации Book Restorer (BKR Automation) позволяет импортировать, редактировать и сохранять страницы книг в пакетном режиме. (Режим отложенной обработки.) Это удобно при наличии огромного количества страниц похожего вида с одинаковыми принципами обработки, а также при обработке нескольких многостраничных книг. Сначала осуществление подготовки заданий для нескольких книг, а только затем пакетная обработка этих книг без участия оператора.

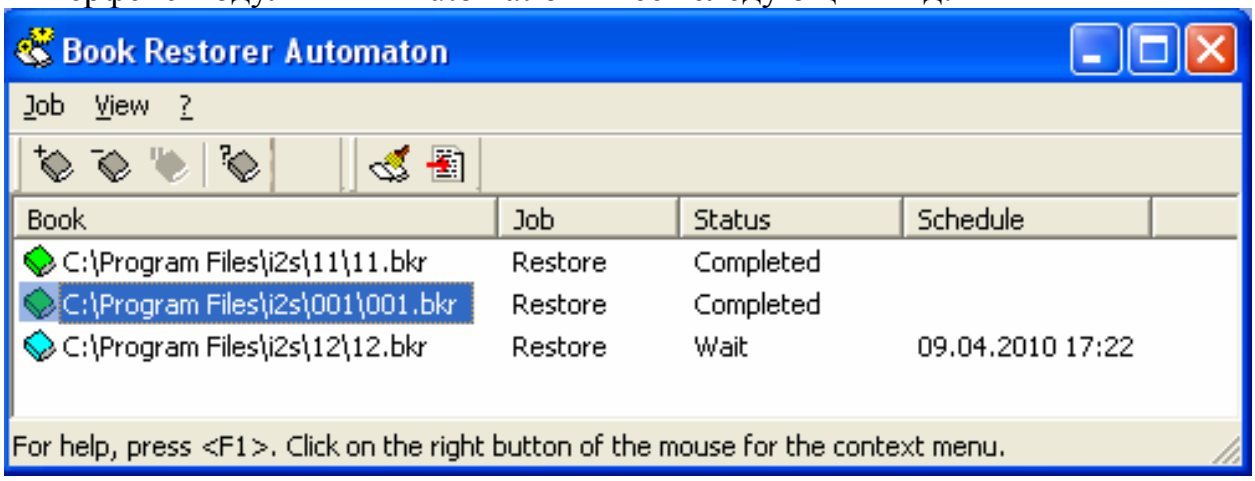
Сложное задание (различные процессы для каждой страницы)



Однородное задание

То есть модуль автоматизации позволяет либо применить готовые скрипты (соответствующим образом сформированный набор процедур обработки) к необработанным однотипным сканам, либо для сложной книги (точнее, книг) с разными страницами сформировать задание, которое выполнить позже.

Интерфейс модуля BKR Automation имеет следующий вид:



На панели находятся две группы кнопок. Первая из четырех:

- открыть задание

- удалить задание

- прервать текущее задание

- показать свойства задания

Вторая - из двух:

- открыть текущее задание (книгу) в Book Restorer

- экспорт журнала (лога)

Дальше идет таблица с перечнем заданий. Каждая строка таблицы содержит иконку статуса задания; имя книги; вид задания (импортировать/обработать/напечатать); статус; запланированное время выполнения.

Задание может иметь один из следующих статусов:

- ожидание

- ожидание разблокирования

- прервано (interrupted)

- в обработке

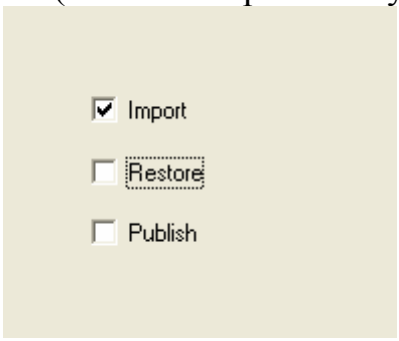
- выполнено

- ошибка

Книга, предназначенная для обработки, не должна быть открыта в Book Restorer для редактирования. Это в случае необходимости можно проверить с помощью кнопки открыть задание в Book Restorer (заблокирована ли книга).

Использование модуля автоматизации предполагает следующую последовательность действий:

1) Кликнуть на иконку или в меню job выбрать команду add, чтобы добавить задание (книгу). Будет предложено выбрать из трех вариантов заданий (можно выбрать любую комбинацию):

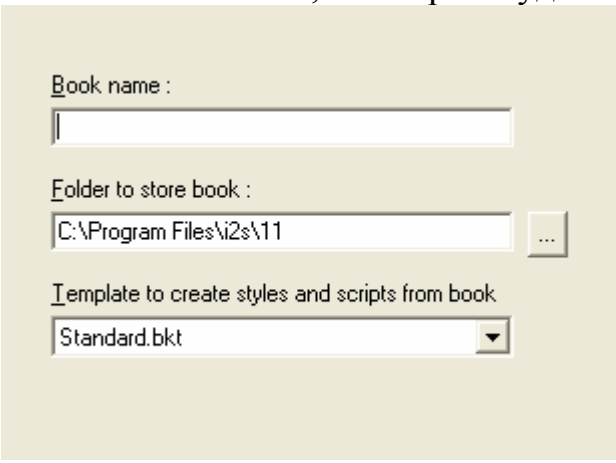


- импортировать;
- обрабатывать;
- опубликовать.

Импортировать (import) - необработанные сканы будут импортироваться в папку книги. Саму книгу нужно будет создать.

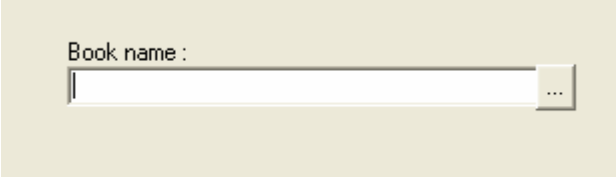
Обрабатывать (restore) - обработанные изображения будут храниться в виде временных файлов в папке книги. Публиковать (publish) - помимо сохранения результатов в папке книги, изображения итоговых страниц будут в заданном формате и разрешении сохранены в назначенной папке. Поэтому при выборе этого задания следует указать путь к целевой папке.

2) Если в том числе выбрана опция «импортировать», то откроется диалоговое окно «New», в котором будет создана новая книга:

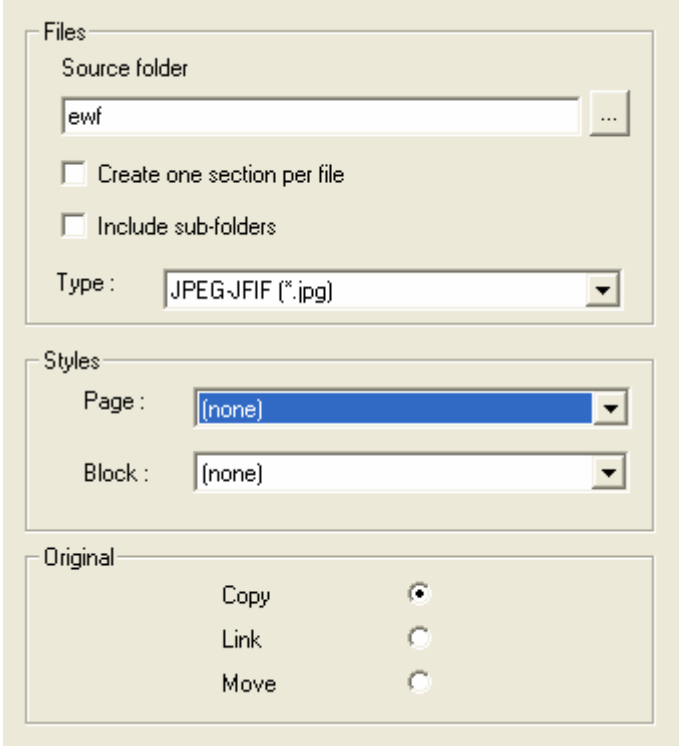


В диалоговом окне следует указать имя книги, папку, в которой будет сохранена книга, а также файл с шаблонами стилей и скриптов для книги.

Если опция «импортировать не выбрана, то потребуется указать, какую уже существующую книгу следует добавить в обработку:



3) Процедура импортирования. Если было выбрано «импортировать», то откроется следующее диалоговое окно:

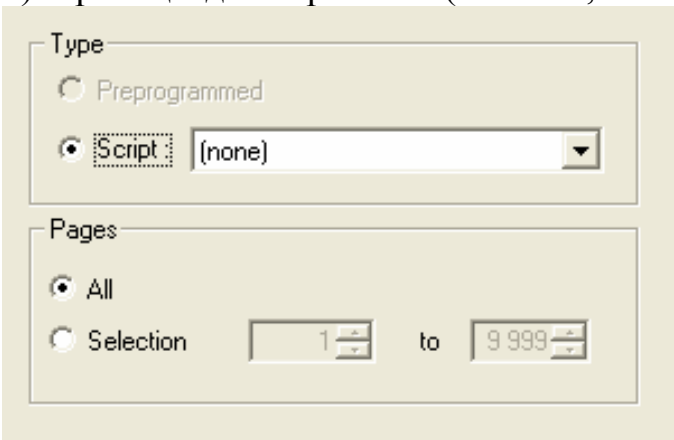


- указать папку, в которой лежат изображения, предназначенные для импорта
- поставьте галочку «create one section per file» - одна секция на файл, чтобы для каждого импортируемого файла создавалась новая секция (имеющая имя импортируемого файла)
- поставьте галочку recursive, если нужно, чтобы файлы импортировались и из подкаталогов указанной папки.
- выберите тип импортируемого файла (tif, jpg, bmp, pdf, png)
- выберите стиль блока и стиль страницы
- в зоне ORIGINALS укажите, что делать с исходными изображениями (link - использовать оригинальные исходники, move - переместить исходники в папку книги, copy - копировать исходники в папку книги)

4) Процедура обработки. Если на первом шаге было выбрано «restore», то откроется диалоговое окно, в котором следует выбрать:

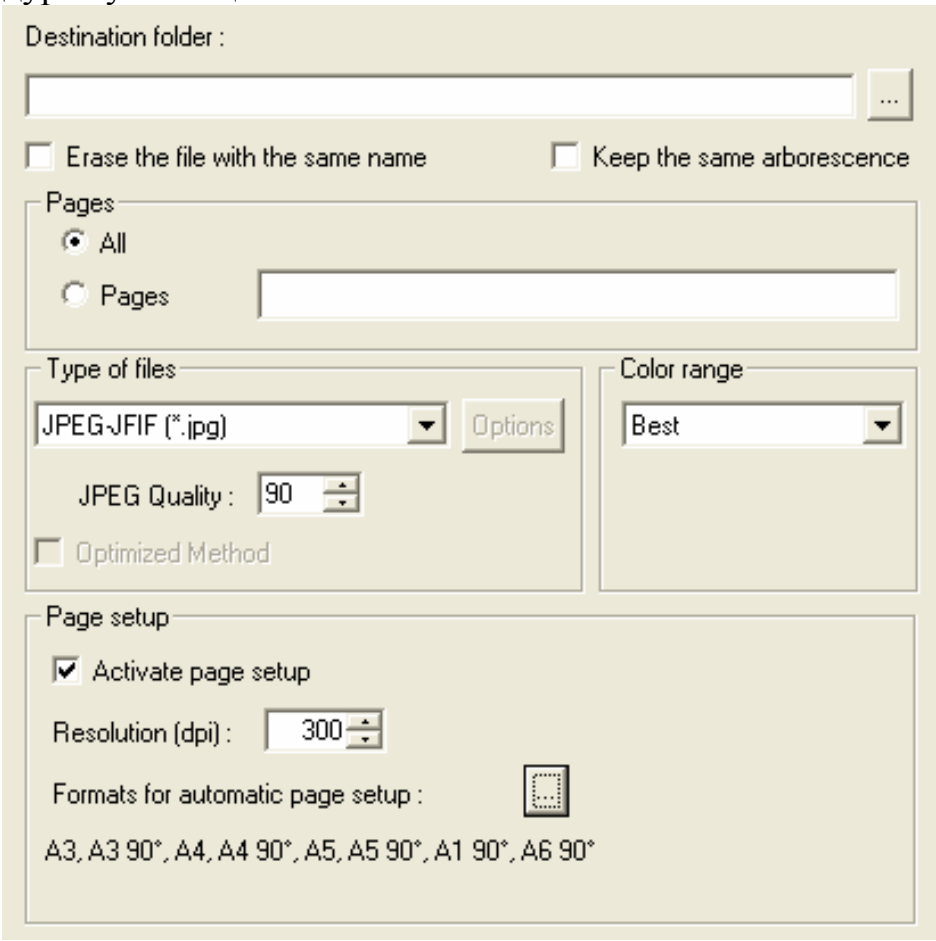
а) тип работы (preprogrammed или скрипт);

б) страницы для обработки (All - все, Selection - диапазон страниц)



Опция «preprogrammed» означает, что книга создана ранее в Book Restorer и уже запрограммировано, какие процедуры и скрипты применяются к изображениям страниц. Опция «Script» предоставляет возможность выбора из скриптов, содержащихся в шаблоне, выбранном на шаге 2.

5) Процедура публикации

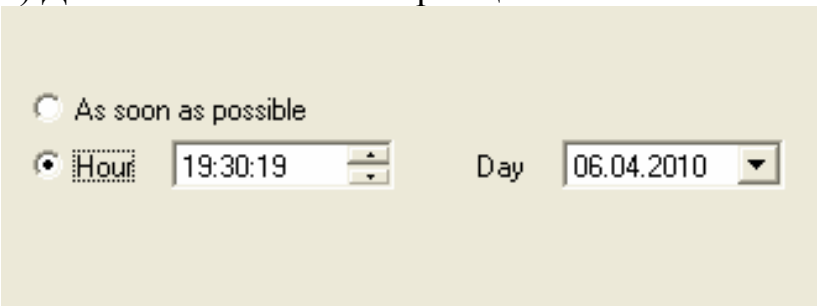


При запуске публикации:
- укажите целевую папку
- стирать/не стирать файлы с уже имеющимися именами
- тип публикуемых файлов

а) для jpg-ов укажите качество 0-100

б) для tif-ов можно выбрать многостраничный тип, а также опцию «один файл на секцию» - тогда будет создан набор многостраничных tiff
- цветность (черно-белый/цветной/серый/автоматический выбор)
- разрешение
- предустановленный формат (размер) страниц

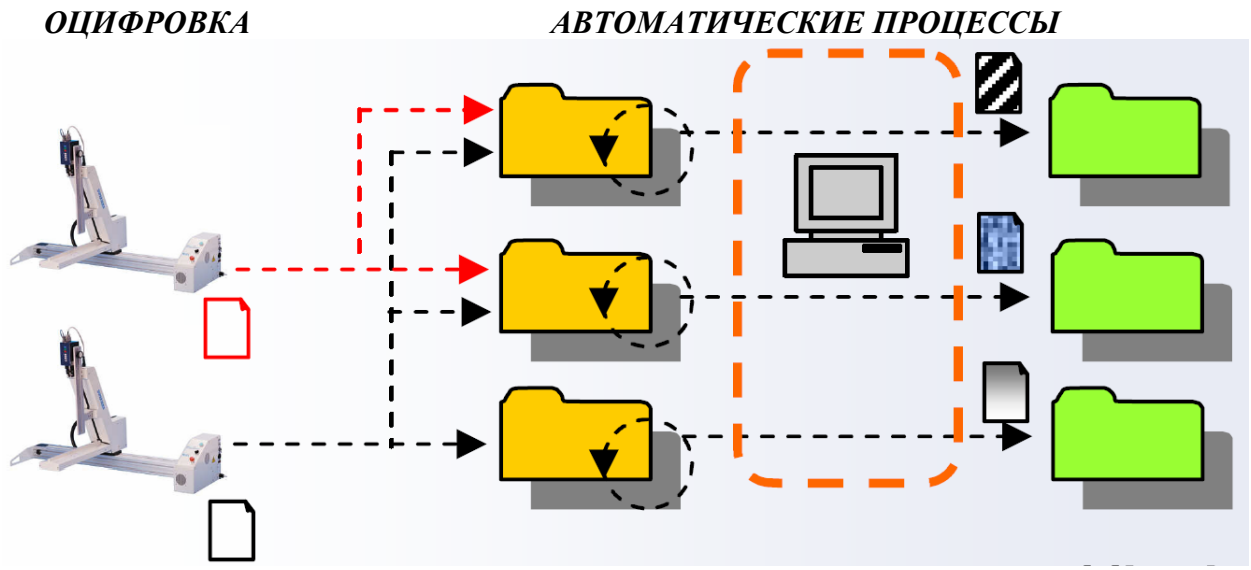
6) Диалоговое окно планировщика «Shedule»



Можно выбрать опцию «как можно быстрее» для запуска обработки сразу же после завершения текущего процесса. Можно также выбрать опцию «Hour» и задать точное время (часы, минуты) начала процесса обработки.

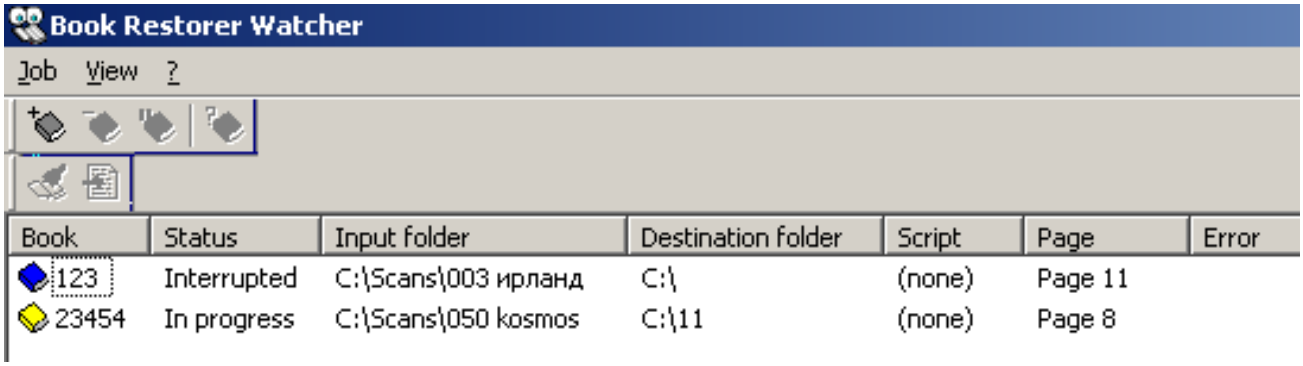
5.2. BKR Watcher - контроль за папкой захвата

Модуль BKR Watcher позволяет автоматически следить за папкой (директорией), в которой сохраняются поступающие с участка сканирования (потокового сканера) необработанные изображения, и обрабатывать их.



Данный модуль улучшает производительность так как обработка (реставрирование) осуществляется одновременно со сканированием

Интерфейс модуля BKR Watcher в целом повторяет интерфейс модуля автоматизации. Такое же точно меню и панели кнопок:

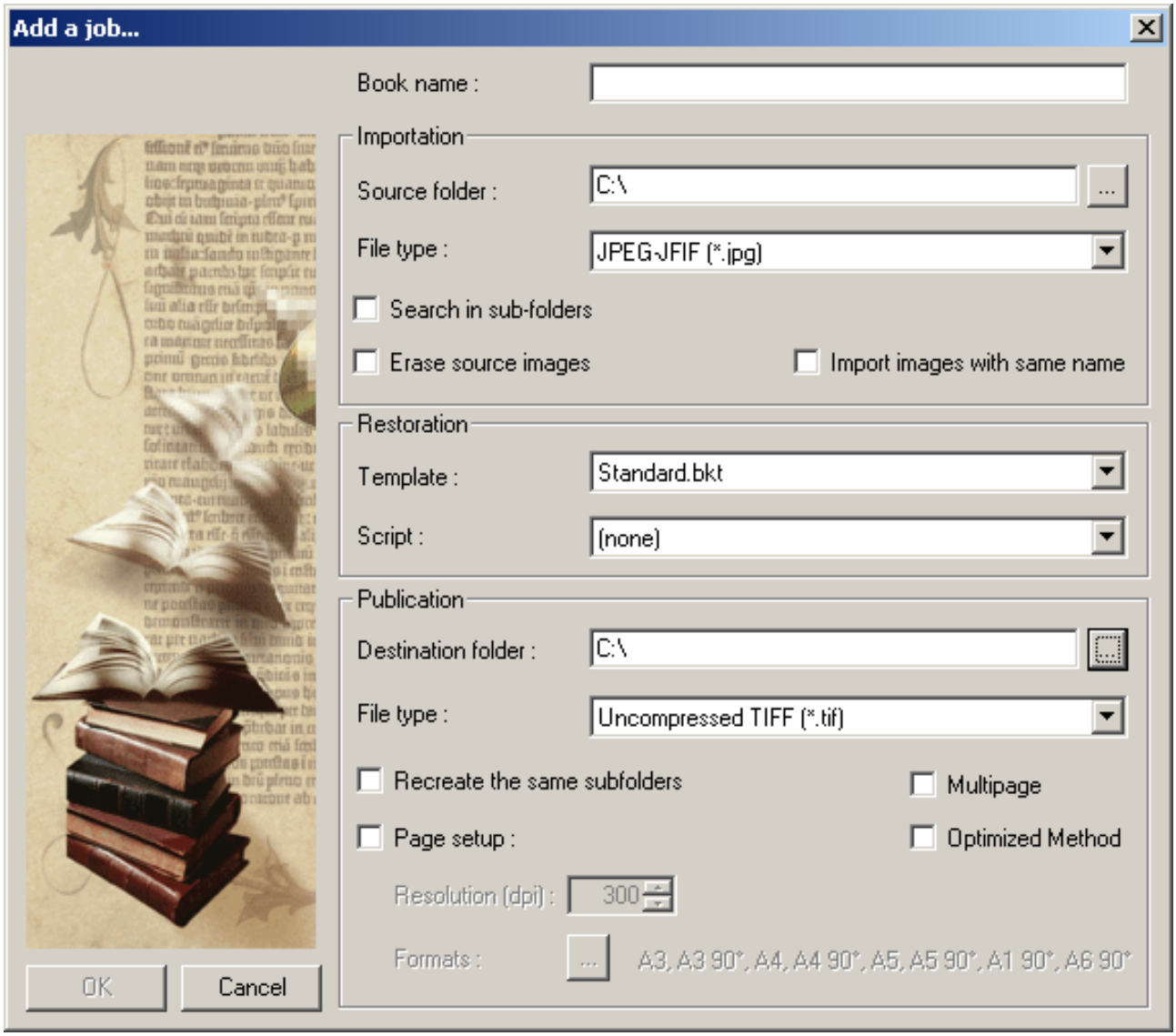


Однако, если полями таблицы в BKR Automate были "книга - задание - статус - время", то в BKR Watcher это " книга - статус - исходная папка - папка назначения - скрипт - страница - ошибки".

В столбце "скрипт" показывается имя выбранного скрипта, который применяется к файлам в исходной папке. Разнородные сканы следует складывать в разные папки, чтобы к ним применялись разные скрипты.

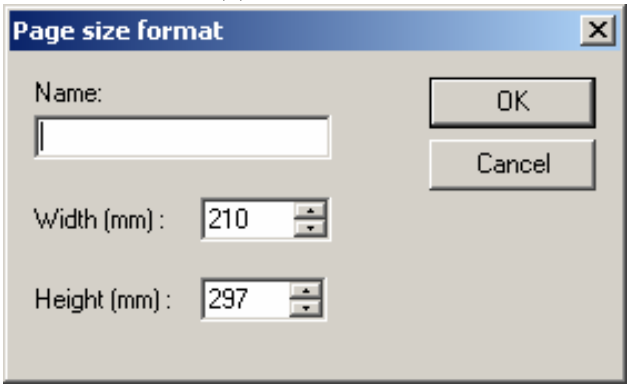
В столбце "страница" указывается текущая обрабатываемая страница, после обработки -последняя.

В отличие от сложного многоступенчатого диалога в модуле BKR Automate здесь все параметры задаются в одном диалоговом окне (форме):



Указывается имя книги, выбирается папка-источник, подлежащий применению к файлам в этой папке скрипт, и целевая папка.

При публикации изображения (сохранении обработанного изображения в целевую папку) можно задать те же параметры "печати", что и в BKR Automate: разрешение, размещение изображения на странице определенного формата A3, A4, и т.д. (который будет служить как бы фоном изображения) или самостоятельно заданного:

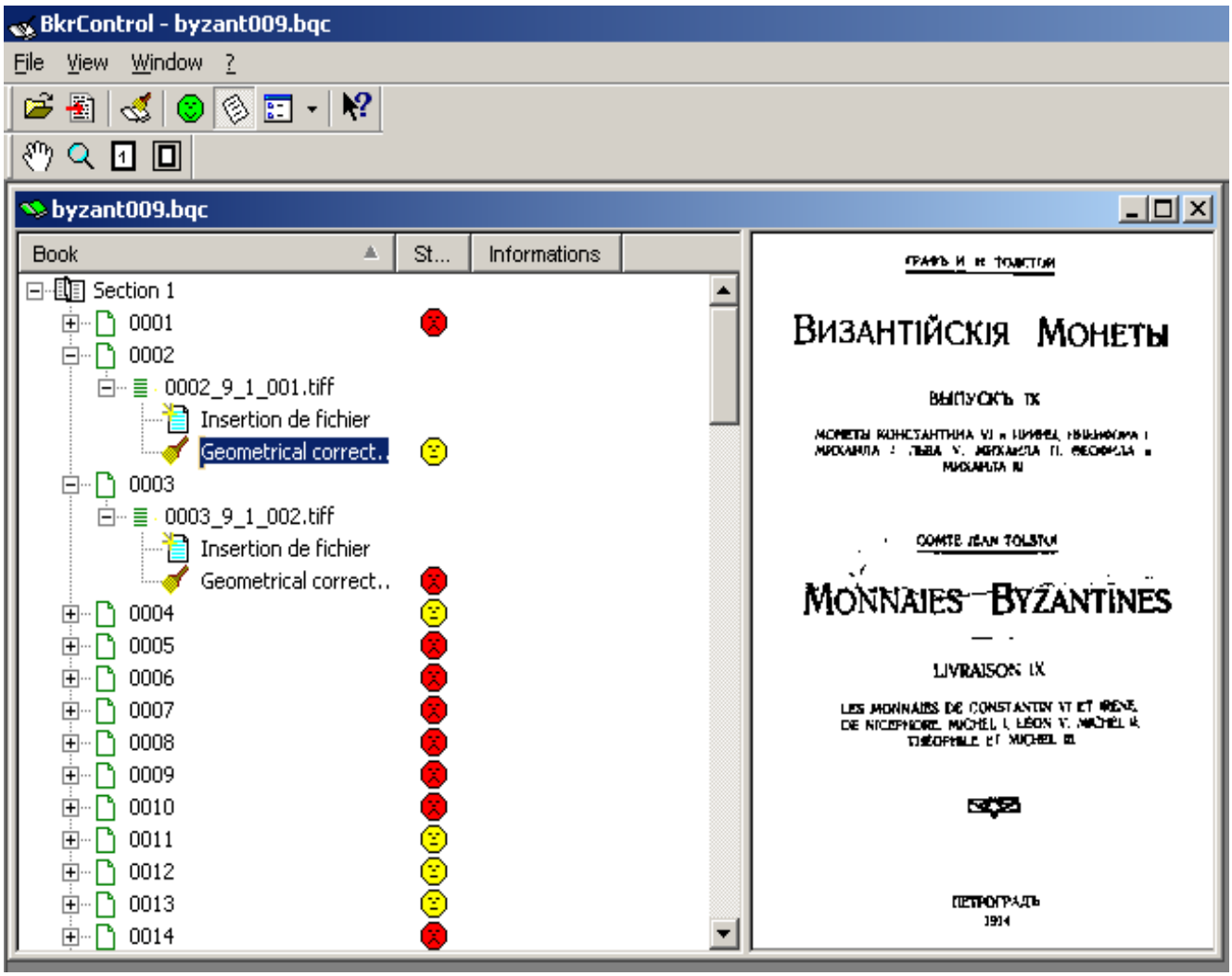


5.3. BKR Control - контроль качества

Модуль контроля качества Book Restorer (BKR Control) используется для подготовки отчета о процессах обработки, применимых к изображениям в книге, созданной Book Restorer.

Для каждого изображения и каждого процесса рассчитывается коэффициент качества, который позволяет определить, какие изображения следует проверить оператору.

Интерфейс программы имеет следующий вид:



Рабочее окно программы включает:

1) две панели инструментов;

2) список процессов со статусами (коэффициентами качества),

3) зону страницы (показывает результат)

Первая панель инструментов содержит следующие кнопки:

открыть документ (файл формата *.bqc)

экспортировать журнал в текстовый файл

откроется диалоговое окно для выбора необходимой информации:

открыть документ (книгу) в программе Book Restorer

показать коэффициент качества для выбранного процесса

показать дополнительную информацию о выбранном процессе

переключиться между списком в виде дерева и списком в виде таблицы

открыть файл справки (показать информацию о команде)

Вторая панель инструментов содержит кнопки для работы с изображением в зоне страницы:

перемещение изображения нажатой левой кнопкой мыши

масштаб (при зажатом ctrl - уменьшение)

реальный размер изображения (масштаб 1:1)

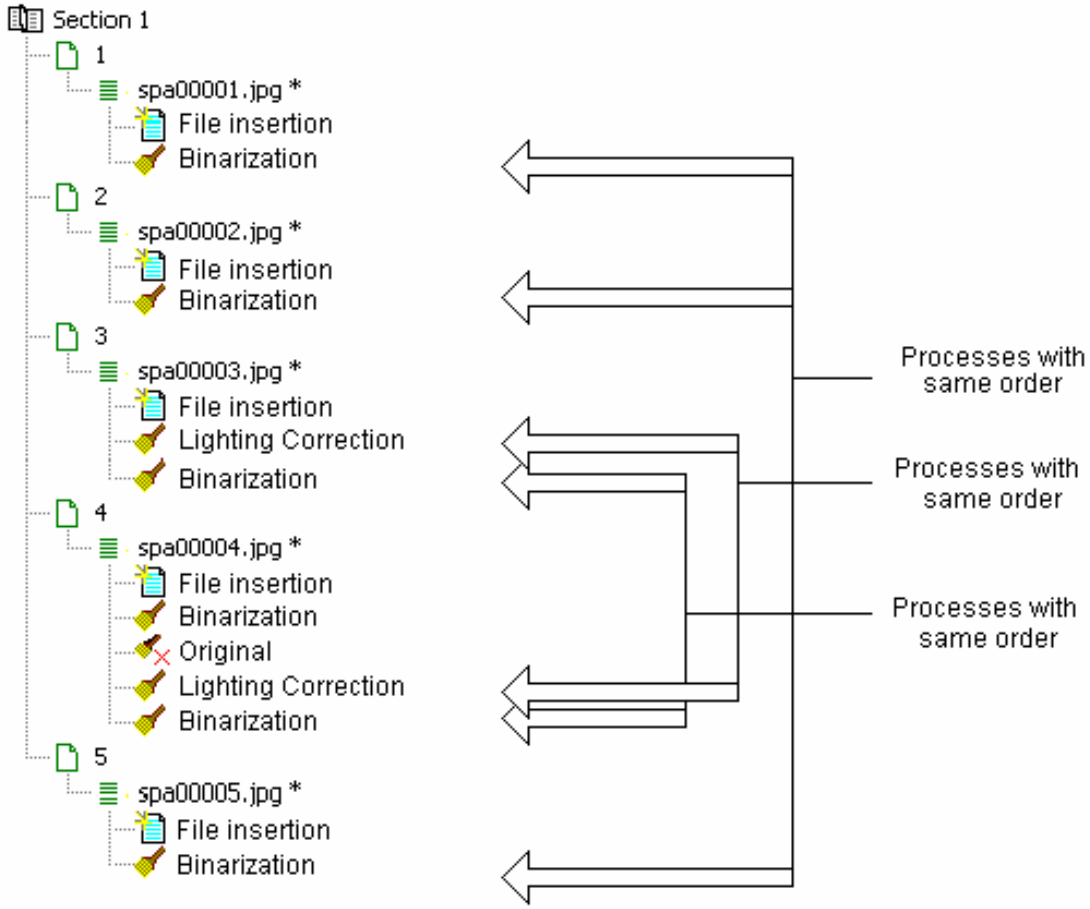
показать выбранный блок в зоне страницы целиком

Список процессов в левой части рабочего окна может быть представлен не только в виде дерева, но и в виде таблицы.

Коэффициент качества для процесса определяется путем проверки результатов применения этого процесса ко всем остальным страницам книги. Чем больше схожих однотипных страниц в книге, тем надежнее рассчитанные коэффициенты. Программа Book restorer на самом деле всегда сохраняет эту информацию в файле *.bqc, просто без соответствующего модуля ее не увидеть. Учитываются следующие процессы:

- изменение гистограммы («Уровни»);
- бинаризация;
- компенсация поворота;
- деспекл;
- коррекция геометрических искажений;
- выравнивание освещенности;
- минимальный прямоугольник (при автоопределении зон).

Каждая секция рассматривается независимо от расчета других компонентов качества:



В меню команда Preferences позволяет выбрать пороговые значения (указывается состояние качества процесса красным, желтым или зеленым цветом), а также период обновления (приведение списка в соответствие с открытой книгой, в которой может идти автоматическая обработка).

5.4. Автоматизация применения Book Restorer средствами OLE

Программа Book Restorer может быть запущена внешним приложением с помощью средств OLE-Automation (таким образом сканы книг могут обрабатываться из приложений пользователя):

Для этого Book Restorer содержит библиотеку типов - файл ***BookRestorer.tlb***, которая содержит информацию о классах, типах и интерфейсах.

Библиотеки типов могут быть просмотрены с использованием различных средств, например таких как Microsoft OLE/COM Object Viewer (oleview.exe, входит в Microsoft Platform SDK) или Object Browser в Visual Basic (до версии 6), а также Visual Studio.NET.

Для этого Book Restorer нужно запустить один раз, чтобы исполняемые файлы были прописаны в файле реестра, а затем подключить библиотеку типов.

Простейшее собственное приложение для обработки можно создать с использованием Visual Basic в Excel (функции любого приложения, написанного с использованием OLE-интерфейса можно вызывать из Excel). Excel при запуске читает раздел библиотек типов реестра, компилирует список всех OLE-библиотек и загружает библиотеки Visual Basic и Excel.

А) Подключение OLE-библиотеки BookRestorer:

1. Откройте VBA Editor (*Tools\Macros\Visual Basic Editor*)
2. Кликните на Browse и укажите путь к файлу BookRestorer.tlb
3. Выберите Tools/References и поставьте галочку у Book Restorer

Б) Просмотр свойств и методов, функций и переменных Book Restorer

1. Откройте VBA Editor (*Tools\Macros\Visual Basic Editor*)
2. Выберите команду **View / Object Explorer**

В) Объявление переменной объект книга "Book Restorer":

Dim Book As BookRestorer.Document

Г) Создание в Visual Basic формы с кнопками для вызова функций Book Restorer

Как уже упоминалось, доступные файлы справки к Book Restorer 3.1. включают файл AutomationOLE.hlp, содержащий информацию о доступных функциях и переменных. Использование средств OLE позволяет комбинировать обработку в Book Restorer с другими программами.

Доступные функции

объект	функция	тип	синтаксис	действие функции	параметры	возвращаемый результат
книга	Open	void	*Open(String Filename)*	открывает существующую книгу	*Filename*: имя и путь к существующей книге	
	Close	void	*Close()*	закрывает открытую книгу		
	Save	void	*Save()*	сохраняет открытую книгу (отличается от функции «опубликовать»)		
	NewBook	void	*NewBook(String Name, String Folder, String TemplateFilename)*	создает новую книгу	*Name*: имя создаваемой книги, *Folder*: папка, в которой будет создана книга, *TemplateFilename*: имя и путь к шаблону книги (может быть пустым)	
	Import	long	*Import(String FileName, long AfterPage)*	импортирует один или несколько файлов с изображениями (tif, jpg, bmp, pdf)	*FileName:* имя и путь к импортируемому файлу, *AfterPage:* номер предыдущей страницы (0, чтобы вставить в начало книги)	импортированная страница
	AddPageSize	void	*AddPageSize(String lpszName, double dblWidth, double dblHeight, bool Enable)*	добавляет формат (стиль) страницы	*LpszName:* имя формата (стиль страницы), *DblWidth:* ширина в миллиметрах, *DblHeight:* высота в миллиметрах, *Enable:* разрешить формат (да/нет)	

страница	NewPage	long	*NewPage(long AfterPage)*	добавляет пустую страницу в книгу	*AfterPage:* номер предыдущей страницы (0, чтобы вставить в начало книги)	созданная страница
	GetPage	long	*GetPage(long Index)*	возвращает указатель на заданную страницу	*Index:* индекс предыдущей страницы (0 для первой страницы)	импортированная страница
	GetCaption	String	*GetCaption(long Object)*	возвращает имя объекта (страницы)		строка с именем указанного объекта
	DeletePage	void	*DeletePage(long Index)*	удаляет страницу в книге	*Index:* индекс страницы к удалению (0 для первой страницы)	
	SelectPage	void	*SelectPage(long Index)*	выделяет страницу в книге	*Index:* индекс выбранной страницы (0 для первой страницы)	
стиль	GetStyle	long	*GetStyle(String Caption)*	возвращает указатель на заданный стиль	*Caption:* имя стиля	указатель на стиль
процесс обработки	GetScript	long	*GetScript(String Caption)*	возвращает указатель на заданный скрипт	*Caption:* имя скрипта	указатель на скрипт
	Restore	void	*Restore(long Page, long Script)*	применяет скрипт к странице	*Page:* страница, к которой должен быть применен скрипт, *Script:* указатель на скрипт (результат функции GetScript)	
публикация	Publish	void	*Publish(long Page, String Folder)*	печатает заданную страницу	*Page:* указатель на страницу к публикации (результат функции GetPage), *Folder:* путь к папке для сохранения результирующего изображения	